DRAFT

IMPROVING ACCURACY AND REDUCING COSTS OF ENVIRONMENTAL BENEFIT ASSESSMENTS

Risk Communication for Superfund Sites An Analysis of Problems and Objectives

by

Gary H. McClelland
William D. Schulze
Don L. Coursey
Brian Hurd
Julie R. Irwin
Rebecca R. Boyce

USEPA COOPERATIVE AGREEMENT #CR812054-02

September, 1987

PROJECT OFFICER

Dr. Alan Carlin
Office of Policy, Planning and Evaluation
U.S. Environmental Protection Agency
Washington, D.C. 20460

DISCLAIMER

Although prepared with EPA funding, this report has neither been reviewed nor approved by the U.S. Environmental Protection Agency for publication as an EPA report. The contents do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

TABLE OF CONTENTS

RISK COMMUNICATION FOR SUPERFUND SITES: AN ANALYSIS OF PROBLEMS AND OBJECTIVES

RISK COMMUNICATION FOR SUPERFUND SITES: AN ANALYSIS OF PROBLEMS AND OBJECTIVES

CHAPTER 1: INTRODUCTION AND OVERVIEW

1.1 The Problem

The Environmental Protection Agency has an extraordinarily difficult task in managing the Superfund Program. Although citizens are very concerned about Superfund sites, scientific estimates of the risks from most sites indicate that the hazards are very small. Thus, EPA faces public demands for extensive and expensive clean-up for many sites when scientific risk assessment fails to justify such efforts. Given these circumstances as well as the real fears of residents living near Superfund sites, the hypothesis underlying most research on risk communication has been that people just do not understand the "true" risks and that good communication of these risks will in and of itself solve or reduce the problem defined above. In other words, good risk communication is just good communication. This viewpoint, which is not supported in the research reported here, comes in great part from experiences with natural hazards such as tornadoes, hurricanes and floods where risk communication programs have shown considerable success. Sorensen and Mileti (1987), for example, argue that the social psychological model developed for risk communication about natural hazards provides many insights into good risk communication. We show in the research presented here that the argument that risk communication is just good communication is likely false for low probability risks such as those present at most Superfund sites. In fact, when the standard risk communication strategies from natural hazards research (which work well at high probabilities as Sorensen and Mileti suggest) were applied in a low probability situation, that of possible volcanic activity near the Mammoth Ski area in California, an utter fiasco resulted. Property values fell substantially after a warning was issued, even though the odds of a Mt. St. Helens type event were minute. Further, these odds were clearly described to residents and the news media (See Thayer et al., 1986). We have previously documented a similar substantial fall in property values near a Superfund site where scientific assessments have shown no evidence of significant risk (Schulze, McClelland and Hurd, 1986). Thus, both Superfund sites and natural hazards with low probabilities of harmful consequences have been shown to have the same potential for inappropriate levels of concern among populations exposed to or hearing about the risk. Consequently we agree with Slovic's (1986) assessment that, for low probabilities, we know almost nothing about risk communication.

The primary purpose of the research which we present here has been to attempt to determine what goes wrong with risk communication at low probabilities. The research initially attempts to answer two related questions. First, does something go wrong in the way people think about low probability hazards? Second, can it be conclusively shown that the individual and community response to low probabilities is inappropriate? The answers we provide to these two questions in turn raise some very serious policy issues for EPA with respect to risk communication at Superfund sites (which we discuss later in this chapter). Since we find that people do fail to behave appropriately at low probabilities <u>no matter how well they understand the risk,</u> we then analyze how factors that affect risk judgments may be employed as part of a risk communication strategy to help people better judge the risk from Superfund sites. That strategy in our view must incorporate an understanding of the inherent inability of the public to respond appropriately to even the clearest most understandable statements concerning low probability risks.

1.2 Summary and Implications of the Research

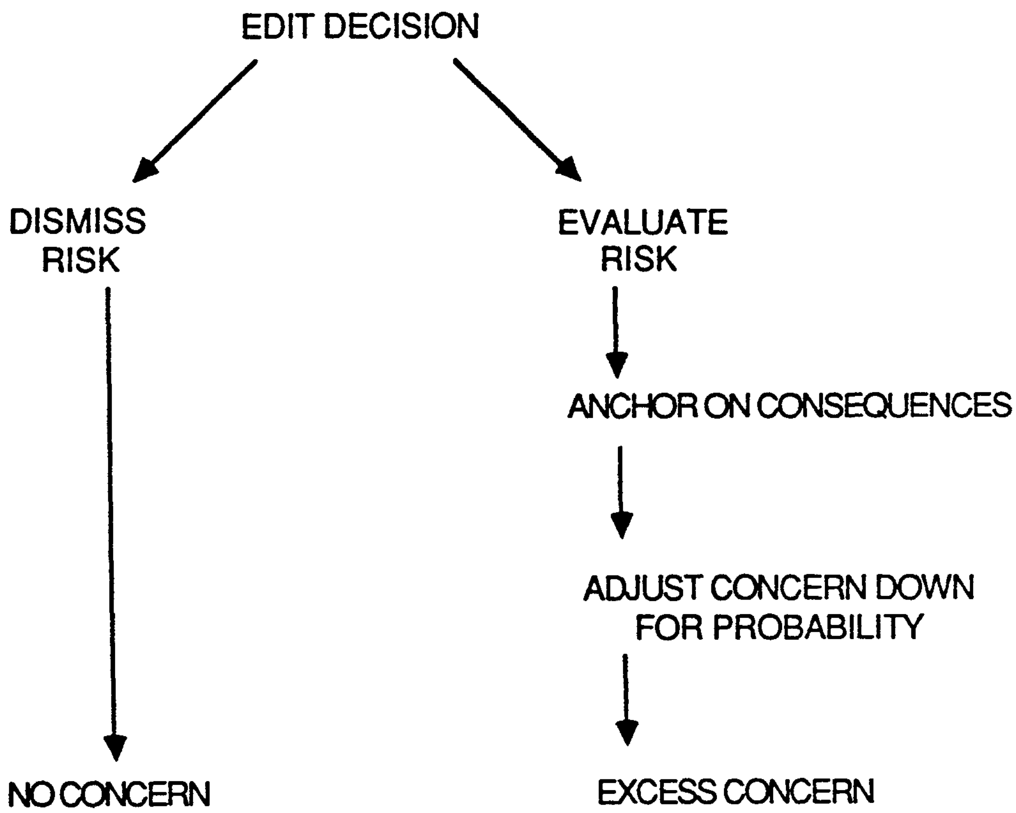
This section summarizes two research efforts used to develop specific suggestions for risk communication at Superfund sites. The first effort sought to expose people to a series of situations where the risks were made absolutely clear so the response to perfect risk communication could be examined. The probability was experimentally lowered to see what difference low as opposed to high probabilities made in the response and to gather data in a simple situation to help understand the source of any problems which might appear at low probabilities. The response chosen to measure concern for the risk in the experiment was to find out how much people would pay for insurance against a risk of loss. The loss was a fixed dollar amount. The probability of loss was obtained by putting red and white poker chips in a bag where subjects were first shown the chips and repeatedly told the number of each color before they saw them placed in the bag. Thus, people knew the exact dollar loss and exact number of red and white chips in the bag. They were then told if a red chip were drawn they would lose the specified amount unless they bought insurance which would prevent the loss if a red chip were drawn. Thus, they could buy their way out of being exposed to the risk, and their willingness to pay was a clear measure of their concern. The analogy to a hazardous waste site is as follows: the fraction of red chips in the bag corresponds to the odds of cancer or other health problems from the site, while the specified dollar loss corresponds to the consequences of cancer or other health problems if they occurred. The point is not that a monetary loss is like cancer or that drawing chips is actually like the risk of getting cancer; but rather, if the mental process generating concern for drawing a red chip fails to yield appropriate behavior at low probabilities in the very simple experimental situation, in the more complex and emotional real world, behavior will be much worse.

In our experiments, the value that people should place on insurance, at least to a first approximation, is the probability times the loss.[1] This amount is called the expected value of the loss. However, based on debriefing of subjects, none of them used this procedure to arrive at the value of insurance. Rather, as Hammond, et al., (1987) have noted, when people do not have access to an analytical mental process (calculation of expected value) to decide how concerned they should be, they use an intuitive mental process instead. In our experiments this intuitive process led to the following results. At higher probabilities of loss, such as four in ten, people intuitively valued and paid expected value for insurance. In other words, at higher probabilities the intuitive mental process yielded about the right level of concern for the loss. However, as we dropped the probability below one in ten the response began to split into two types. Either (1) the level of concern did not fall enough as the probability fell and people paid too much for insurance (many times expected value) or (2) people showed no concern at all, dismissed the risk, and bid zero. Figure 1.1 shows a model of how people judge how concerned to be at low probabilities based both on our debriefing of

[1]We demonstrate that risk aversion plays no role in our experiments in Chapter 3. However, in the real world situation of a Superfund site, risk aversion or risk seeking behavior may play a significant role. Risk preferences, however, have nothing to do with the cognitive errors which occur at low probabilities. These errors are not only clearly demonstrated in the research, described here but have been repeatedly demonstrated by psychologists. See for example research by Kahneman and Tversky, 1979, 1981, 1984, Lichtenstein, et al. 1978, and Combs and Slovic 1979.

FIGURE 1.1

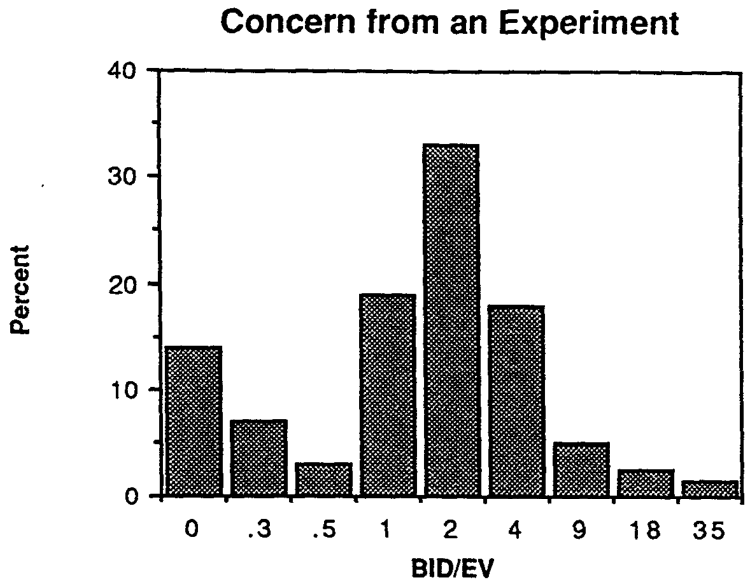
A MODEL OF RISK JUDGMENT

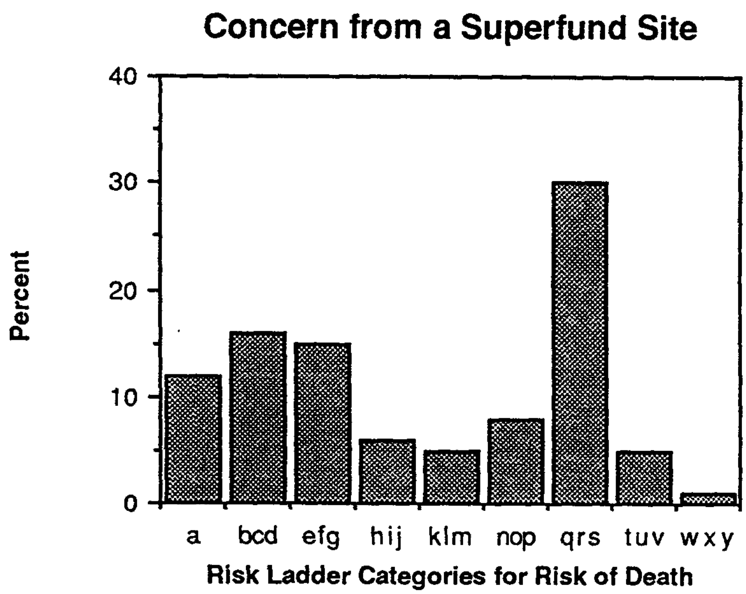


subjects and on the responses described above. Apparently when faced with a low probability people first try to decide if the risk is worth worrying about at all. People face innumerable small risks and evaluating all of them would be impossible. Thus, they first decide whether to evaluate the risk or just to dismiss it. If they edit the risk and dismiss it then they act as if they show no concern. On the other hand if they decide to evaluate the risk they then appear to go through the following process: first, they think about or <u>anchor</u> on the loss event. In the case of a Superfund site the loss event is likely to be cancer, a birth defect or other illness or disease possibly leading to death. In our experiments the consequence is just the loss of a sum of money. In either case people then take this level of concern, obviously large for a Superfund site, and attempt to <u>adjust</u> their concern downward to account for the fact that the consequence is not a certainty. People start out by thinking "wouldn't it be terrible if the site gave me or my family cancer"; and then think, "but maybe it won't, so I don't have to be quite so worried." Unfortunately psychologists have repeatedly demonstrated that when this intuitive anchoring and adjustment thought process is used, the adjustment almost always falls short. In other words, for low probability risks people do not adjust down enough and end up with excess concern. In the case of Superfund sites, since the potential consequences are so bad and the probability is so low, a lot of downward adjustment in concern is necessary. By falling short in the adjustment process people end up much too worried.

One way of showing the response to a low probability risk is to plot the frequency distribution of the amount of concern generated. The upper panel in Figure 1.2 plots the percent of subjects in an experiment offering to pay different amounts for insurance. The amounts shown on the horizontal axis are expressed as amount bid for insurance divided by expected value. Since the

FIGURE 1.2





Annual Risk of Death:
a = no risk
b = one in 9 million
f = one in 100 thousand
j = one in 10 thousand
n = one in one thousand
r = one in one hundred
v = one in ten

loss was $40 and the odds of loss were 1/100 in this experiment, expected value was 1/100 x $40 = $.40. Thus, people should all offer to pay about $.40 for insurance (shown as unity along the horizontal axis). However, the most frequent bid was 2 1/2 times expected value or $1.00 (which falls in the 2 BID/EV category). The second most frequent bid was zero. Note also that the horizontal axis is logarithmic (with the exception of the separate category for zero bids) to compress the scale since some bids were very large. The lower panel in Figure 1.2 shows the frequency distribution of concern for residents around a Superfund site in Monterey Park, California. The vertical axis again measures frequency as a percent of the population while the horizontal axis measures risk beliefs from a risk ladder, again on a logarithmic scale with a separate category for zero risk. The similarity between the two distributions is striking, showing a lower mode of individuals who edit the risk and show little concern, and an upper mode of people who <u>overestimate</u> the appropriate level of concern.

This last statement, that some people overestimate low probability risks, can only be justified on the basis of the experiment, not on the basis of the risk beliefs taken from the hazardous waste site. Who is to say that the upper mode group who believe that the Superfund site is very dangerous will not some day be proven correct, even if scientific assessments now show no risk? In contrast, under the controlled circumstances of the experiment, the true risk is known to all, subjects and experimenter alike. In this situation the fact that subjects behave inappropriately can be clearly demonstrated.

It is interesting to note just how badly subjects performed. We have run subjects at 1/100 odds of a $40 loss through more than 100 trials. Thus, a chip was drawn from a bag containing 99 white and one red chip more than 100 successive times. As noted above, subjects should bid $.40 for insurance on each trial. The logic of this approach is as follows: over 100 draws, one would on average expect one red chip to be drawn for one expected loss of $40. If insurance were purchased by a subject in every trial for $.40, $40 would be spent to avoid an expected $40 loss. In fact, most people in the experiment either obtained insurance for around $1.00 on each trial or bid $.00 and did not get insurance. People who got insurance all the time paid about $100 over one hundred trials to avoid one expected loss over that period of $40. People who always bid zero and never got insurance over one hundred trials typically suffered one loss of $40 but paid nothing for insurance and came out far ahead of those who bought insurance. In fact, over 100 plus trials, those who usually bought insurance often went broke.[2] These subjects left the experiment bewildered, realizing they had done something wrong, but not knowing what. They thought they were “playing it safe" by buying insurance to protect the initial money they were given but were in fact wasting their money and their concern on avoiding a low probability event which did not merit their attention. Therefore it can be seen that, for low probabilities, editing is better strategy when compared to overestimating the appropriate level of concern. This suggests that the actual goal of risk communication in analogous situations in the real world should be to encourage people to edit the risk.

People faced with such risks, in deciding what to do, can be thought of as trying to come up with the appropriate level of concern (expected value) using a defective electronic calculator. They punch in the loss and the numerically small probability. However, the calculator, rather than multiplying the two numbers, either comes back with an answer of zero or an answer which is much too high relative to expected value. At high probabilities (above .1) the calculator works much better, yielding numbers fairly close to expected values.

[2] These experiments used real money supplied by the Council on Research and Creative Work at the University of Colorado-Boulder.

The fact that people are in some sense “lost” at low probabilities was also demonstrated in our experiments. Over many trials the concern people placed on the risk did not remain constant as it should have, since expected value was constant over trials. Rather, when long runs of white draws occurred (inevitable with 99 white and 1 red chip) the level of concern drifted downward. In contrast, after a red chip was drawn, concern would drift upward. Thus, these inappropriate cues influenced the level of concern. Slovic (1986) has extensively analyzed the real world factors which tend, often inappropriately to bias the level of concern. The remainder of this section describes the implications of such influences. Many of these factors, shown in Figure 1.3, play an important role in developing an appropriate risk communication strategy.

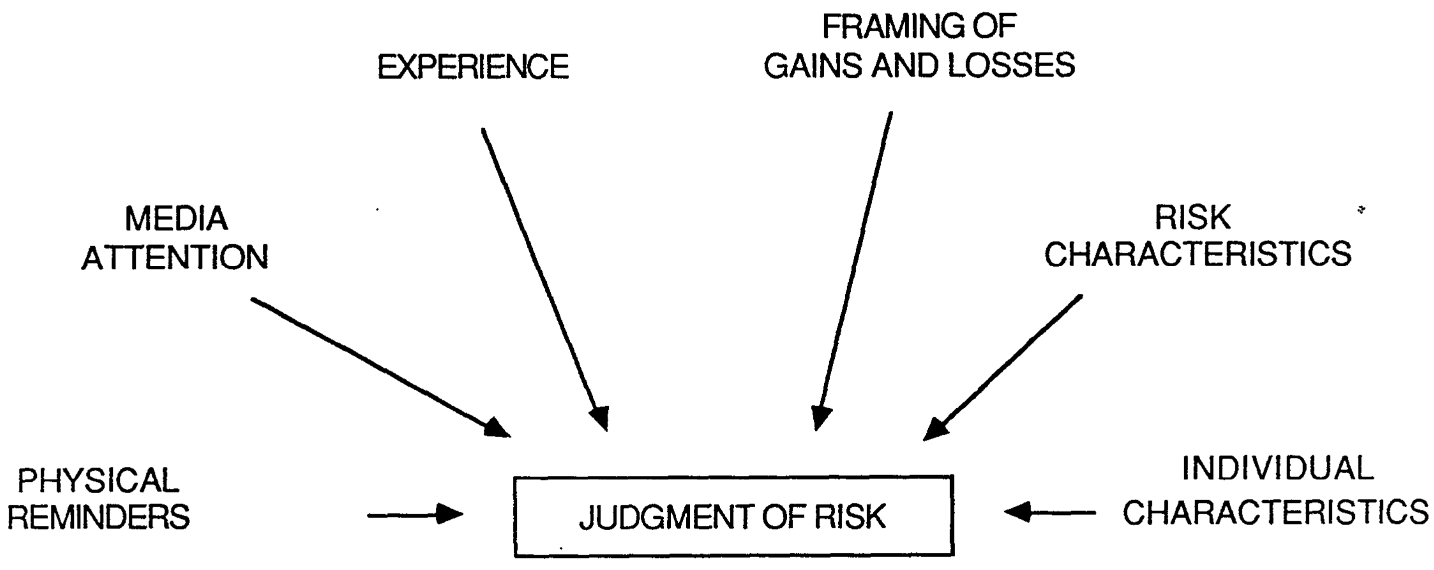
Based on our study of a Superfund site (summarized in Chapter 4), the factors shown in Figure 1.3 work to influence risk beliefs as follows:

Physical Reminders. In the absence of constant reminders people will tend to forget about, or edit, a risk. In contrast, tall fences around a Superfund site, warning signs, the noise of truck traffic associated with clean up of a site, views of workers wearing “space suits,” and odors or smoke from a site all remind people of the existence of risk and tend to cause excess concern. These reminders should be consistent with public safety but should not be exaggerated by a lack of awareness of how the pubic might misinterpret them.

Media Attention. Recent studies have shown that the media contribute substantially to excess concern (see Chapter 5 as well as Wilkins, 1987). Media attention provides reminders of risk when none may actually be present. The media, concerned with ratings or circulation, have an incentive to sensationalize issues especially where public controversy exists. Superfund sites will produce controversy between the no concern

FIGURE 1.3

FACTORS THAT AFFECT JUDGMENT OF LOW PROBABILITY RISK



group and the group with excess concern described above. This controversy will attract media attention which provides reminders of the risk which are inappropriate cues reinforcing the beliefs of the group with excess concern. The media plays a central role in this process which has been termed social amplification. There is little we think EPA can do in this situation other than to avoid contributing to the amplification process by avoiding communicating risks in a way which leads to exaggerated concern.

Experience. Individual experience will strongly influence risk beliefs. If a neighbor has recently died of cancer, that death may well be falsely attributed to a nearby Superfund site by a worried family. If people have been aware of a risk for a very long time, i.e., have a lot of experience with the risk, they tend to adapt to the risk, view it as part of the *status quo*, and edit the risk. Thus, old risks such as coal fired power plants, garden chemicals, and driving are often edited while a new risk such as that from a Superfund site (which used to be viewed as relatively harmless) creates much concern.

Framing of Gains and Losses. People are much more concerned about losses than they are concerned about gains relative to the *status quo*. Thus, reducing an old risk (a gain relative to the *status quo*) is not very important to most people but finding out you have been exposed to a new risk (a loss from the *status quo*) gets people very concerned. Some people may become accustomed to a Superfund site and oppose clean up efforts because any new risk associated with clean up (a loss) increases concern more than removing the old risk from the site (a gain) decreases concern. Risk communication efforts should carefully consider how proposed actions will be stated and interpreted in terms of losses and gains.

Risk Characteristics. Some risks have characteristics which raise concern. For example, people tend to be more concerned if many people are killed at one time. Some modes of death such as cancer may be more feared than others such as accidental death. People may imagine that a Superfund site might explode, spreading a cloud of toxic gas, killing many people immediately, while leaving others to die years hence from cancer. Risk communication about the nature of consequences might do much to allay such fears.

Individual Characteristics. Families with children, younger people, and women all tend to be more fearful of Superfund sites. Education, income level and occupation seem to have little or no impact on risk beliefs. This individual information can be used to help predict where concerns are likely to appear over Superfund sites and over cleanup activities.

Many of the factors discussed above play a role in generating excess concern around Superfund sites. Unfortunately the role of these factors, which have been explored most extensively by Slovic, have been often misunderstood, as has much of Slovic’s research. Slovic is not arguing that these factors are *appropriate* determinants of risk beliefs. Rather, he has demonstrated that a rather long list of mostly irrelevant factors influences the level of concern shown for low probability risks. These irrelevant factors become influential because, as we have argued above, people are lost at low probabilities. Thus, since they have no analytical mental process to rely on, they substitute a very faulty intuitive process which is impacted by the factors shown in Figure 1.3.

Unfortunately, many non-psychologists writing in the area of risk communication, not realizing that the above factors lead to faulty judgments, have incorporated them into suggestions for risk communication. For example, dramatic illustrations which emphasize physical reminders may seem to constitute the elements of good communication, but instead can serve to unnecessarily frighten people. Much current advice on risk communication is ill advised for Superfund sites because the problem in the past has been to get people facing a high probability risk (from tornadoes, smoking, or driving a car) to stop editing the risk and do something (take shelter in a storm cellar, stop smoking, or buckle up, respectively). The problem at Superfund sites is just the opposite, to get the people who are overly concerned to edit the risk and not do anything (do not sell your home at a loss, do not be afraid, etc.).

1.3 Alternative Approaches for Risk Communication.

Given the current understanding of factors which affect risk beliefs, what strategies for dealing with the mismatch between the public’s subjective beliefs and scientific estimates of risk at Superfund sites are available to EPA? Purely for purposes of defining issues and the range of options we briefly evaluate four alternative approaches: (1) Benign neglect, (2) Aggressive risk communication, (3) Conflict resolution, and (4) Complete site cleanup. These alternatives imply (moving from 1 to 4) increasingly expensive cleanup operations.

Benign Neglect. The discussion of the preceding section suggests that Superfund sites, if left completely alone, will eventually become viewed as part of the status quo just like risks from coal fired powerplants, automobiles and garden chemicals. As the new risks from Superfund sites become old risks, concern will subside. Obviously however, for the very reasons outlined above, public concern is now very intense at many sites and would likely take a very long time to subside under this scenario. Beyond the legal and political infeasibility of this approach, it is also bad policy in that people who believe Superfund sites pose great risk are actually suffering. The concern they express is genuine as demonstrated by large drops in property values. These concerns are not feined or insubstantial in spite of our arguments that they result from cognitive problems. Since economic damages are real (as expressed in property values) benefit cost analysis also suggests that action should be taken and that benign neglect is an inappropriate strategy.

Aggressive Risk Communication. Since the intense concern for Superfund sites is of ten inappropriate, one form of action would be to vigorously to manipulate all of the known factors which affect risk beliefs to attempt to decrease that concern and consequently avoid "unnecessary" expenditures for cleaning up sites which generate little actual risk. Two obvious problems are associated with such an approach. First, we do not know enough about risk communication to pursue such a policy successfully. Second, such a manipulative policy is offensive in a democratic society and smacks of policies pursued in closed societies. EPA has an obligation to provide accurate information to the public and, aside from ethical considerations, heavy handed attempts at minimizing risks will inevitably backfire, causing an even more intense over-response from the public.

Conflict Resolution. Since the pubic response to a problem Superfund site will be excess concern among one group of citizens and little or no concern among another group of citizens, conflict will inevitably arise over the appropriate course of action to take in cleaning up the site. This conflict provides the best opportunity for risk communication to provide scientific evidence to both sides in an open public debate. The natural tendency in a formal conflict resolution process (fostered by EPA) will be to reach a compromise. The resulting cleanup effort will likely be substantially less than the very concerned group would have desired. However, the concerned group will recognize, if the process is a fair one, that compromise was necessary to reach a solution given the diversity of public views. EPA should not take sides in this process (avoiding a no win situation) but rather provide technical information through a group of neutral scientists picked by the community to advise them in the conflict. What is critical in this process is that Superfund personnel avoid making the situation worse in the way the site is managed, the way cleanup operations are presented, etc. Chapter 5, which provides our guidelines for risk communication, follows this strategy.

Complete Site Cleanup. In the long run, fairly complete cleanup of Superfund sites, without much consideration of scientific assessments of the risks, in response to public pressure will solve the problem. However, this solution is not without severe disadvantages. First, it is unlikely that enough money could ever become available to satisfy concerned citizens near all Superfund sites. Second, it would obviously be preferable to use risk communication to allow more money to be spent on sites which have higher scientific estimates of risk. Third, cleanup activities themselves tend to frighten the public. The prospect, for example, of burning hazardous wastes on site will be regarded as a new risk and a loss from the current status quo. Both of these characteristics will generate intense concern. Thus, in the absence of a strategy such as conflict resolution cleanup efforts themselves may exacerbate problems.

It is our view that the only practical alternative available to EPA is to apply risk communication in the context of conflict resolution. In many respects, current EPA policy is consistent with this approach. However, it is our view that more effort should be devoted to conflict resolution within the. Superfund program. Current procedures do not provide an EPA sponsored framework for a formal process which includes, for example, the appointment of a community selected panel of scientists to review risks. An example of how risk communication blends with conflict resolution is described in a paper by Hammond, et al. (1976) which is included as the Appendix to this report.

In summary, we are very pessimistic about prospects for explicit risk communication at Superfund sites. If EPA inherits a site where people are not already concerned, attempts at risk communication may well provide inappropriate reminders and cause some people to move from an edit decision to become overly concerned. If EPA inherits a site where many people are already very concerned, attempts at risk communication can easily reinforce those concerns and falsely convince those who currently ignore the site to evaluate the risk. Since most people (experts included) when evaluating a low probability risk tend intuitively to overestimate how concerned to be, they are better off ignoring the risk. Implicit risk communication through management of perceptual cues such as physical reminders may, however, be a fruitful approach along with conflict resolution. Unfortunately most of what we know about explicit risk communication has come from attempts to warn people about important hazards which they are ignoring and so is inappropriate for Superfund sites. Consequently much of the advice on risk communication which EPA has received for these sites is, in our view, incorrect.

1.4 The Organization of the Report.

Readers who do not wish to examine the technical aspects of the research reported here may skip directly to Chapter 5, "Risk Communication Guidelines," which contains our specific recommendations for risk communication procedures at Superfund sites. Chapter 2 provides a technical synthesis of what is known about risk communication and the formation of low probability risk beliefs based both on our own research and the research of others. Chapter 3 presents the details of the core experiments described above. Chapter 4 summarizes our prior study of how risk beliefs were formed at a Superfund site.

CHAPTER 2: A TECHNICAL SYNTHESIS

In this chapter we review what is known about risk beliefs and risk communication from the research literature and especially from our own prior research in this field. This review provides necessary background for our recommendations which are presented in Chapter 5. Based both on the literature and our research we develop a model of risk judgments that leads directly to those recommendations.

2.1 Low Probabilities.

It is quite clear that it is low probability events that are especially problematic for risk judgment and communication. For example, Kahneman and Tversky's research (1979, 1984, and Tversky and Kahneman, 1981) supporting their Prospect Theory shows that outcomes with low probabilities receive a disproportional weight in the decision process. Lichtenstein, Slovic, Fischhoff, Layman and Combs (1978) and Combs & Slovic (1979) have found that many people seriously overestimate low probability events that receive disproportional coverage in the media (e.g., botulism poisoning).

Wallsten and his colleagues (see Wallsten, 1986) provide an interesting demonstration of the difficulties in communicating low probabilities. In one of their experiments, one member of a pair is shown a probability graphically (as an area of a pie chart). That person's task is to communicate that probability to the other member of the pair <u>without using numbers</u>. Instead of numbers the communicator can use phrases such as "very unlikely" or "rarely" or "probably." The person receiving the information then must estimate the probability being communicated. The result is that moderate probabilities (.15 to .85) are communicated rather accurately but more extreme probabilities (close to zero or one) are communicated quite poorly. The error is that the receiver of the communication overestimates low probabilities and underestimates high probabilities. These results suggest that non-numerical language for communicating very low and very high probabilities may simply not exist in English.

Our own laboratory research also dramatically illustrates risk judgment problems for low probabilities. In our studies, participants bid for insurance to protect them against a real economic loss that would occur with a known probability. We used an auction procedure (a multi-unit Vickery or competitive auction) known to have demand-revealing properties in order to get participants to bid what the insurance against the risk was really worth to them. Thus, their insurance bids can be considered a measure of their concern about the potential hazard. Figure 3.1 of Chapter 3 shows the mean values of the ratio of the bids (B) to the expected value (EV which equals probability of the loss times the magnitude of the loss) as a function of probability. If bids are consistent with standard assumptions of economic rationality, then B/EV should equal 1 for all probabilities. For probabilities of .2 and above, B/EV is very close to 1; however, B/EV is somewhat greater for $p = .1$ and very much so for $p = .01$ where insurance bids are about 2.5 times greater than EV. Thus, average concern for low probabilities is much greater relative to an EV model, than average concern for higher probabilities.

Camerer and Kunreuther (1987) report that in their experiments using double-oral auction markets insurance prices approach expected value for a large range of probabilities and loss amounts. They explicitly express surprise at our results reported in Chapter 3 because they obtain mean bids near expected value even for low probabilities. The explanation for the difference in the results between the two sets of studies is that

Camerer and Kunreuther recruited students from decision sciences and finance classes at the Wharton School and Camerer and Kunreuther themselves note that these students “were familiar with concepts of probability, expected value, and sometimes risk-aversion." While it is comforting to know that students explicitly taught expected value learn it and can apply it in these bidding experiments, it provides no evidence that typical citizens can understand and appropriately respond to low probability risks.

Thus, research is consistent in suggesting that low probabilities will be difficult for typical people to understand. By low, we mean probabilities of about .05 and below. Of course, many of the risk probabilities that are of interest to the Environmental Protection Agency, in general and to Super fund, in particular, are much less than that. It is not unreasonable to extrapolate from our data and expect that average concern about a risk will deviate even further from EV for very low probabilities such as 10^{-6}. It should be noted that we abstract from risk aversion in this discussion for clarity.

Figure 3.1 in Chapter 3 suggests that high probability risks (i.e., probabilities >.1) are likely to be well-understood. This is consistent with experience in the natural hazards area. While it is difficult to get people to plan appropriately for the 100-year flood which has a probability of .01, it is relatively easy to get people to respond to warnings of a higher probability such as alerts for imminent floods, tornadoes, or hurricanes. (Although in such cases it is sometimes difficult to communicate to people the appropriate action they should take.)

In summary, the problems of understanding and communicating low probability risks will be very different from those of high probability risks. It is therefore important to make this distinction in any discussion of risk judgments.

2.2 Bimodality.

The obvious next question is what goes wrong in judgments of risks at low probabilities. We believe our research is unique in providing explanations of the problems of judgments of low probability risks. Our explanations are based on natural cognitive processes. We were led to these explanations by discovering that the aggregate picture of Figure 3.1 is very misleading. Figure 3.2 shows the frequency distribution of B/EV for three probabilities (p = .01, .2, and .9) from our laboratory experiments. For p = .9, B/EV has approximately a normal distribution (note that the x-axis is a log scale) centered about 1, consistent with an EV model. The picture is similar for p = .2 although the distribution is more log-normal. However, the picture is very different for p = .01 where there is clear indication of bimodality. A sizeable proportion of people (about 25%) dismiss the risk out of hand (as indicated by bidding zero for the insurance) while another group of people bid at or substantially above EV. Thus, the high mean for B/EV for p = .01 in Figure 3.1 obscures the fact that many people are unconcerned about the risk. Somewhat paradoxically, for low probabilities more people are unconcerned about the risk, but those who do not dismiss the risk are much more concerned, relative to EV.

We think the apparent paradox can be explained by two cognitive processes: editing and anchoring and adjustment. Each individual is confronted by so many low probability risks that it would be paralyzing if one attempted to decide on an appropriate response to all of them. Thus, a useful strategy is to dismiss or "edit" risks that one considers to be below some threshold. For a fixed risk consequence (like the loss in our lab experiments), editing ought to increase as the probability of the risk decreases, or, conversely, the fraction of people concerned enough about the risk to bid for insurance ought to decrease as the probability decreases. For our data, f^+, the fraction of people bidding for insurance is modeled by

$$f^+ = \underset{(152)}{.936} - \underset{(13.2)}{.002}p^{-1}$$

$$DF = 4 \qquad R^2 = .98$$

The data and the model fit are plotted in Figure 3.3. Obviously, the fraction of positive bids falls sharply as the probability falls and the amount of editing increases.

If someone doesn't dismiss the risk and therefore is concerned about the risk, how does that person decide on an appropriate level of concern? In our experiments, deciding on an appropriate level of concern is equivalent to deciding how much to bid for insurance. We think an anchoring and adjustment process explains the mental steps involved in generating a number for a bid. To arrive at a bid, individuals focus or <u>anchor</u> on the loss and then <u>adjust</u> their bid downward to reflect that the loss will occur only some of the time. A consistent result from the cognitive psychology literature (Poulton, 1968; Tversky & Kahneman, 1974; Lichtenstein, et al. 1978) is that such adjustments are almost always underadjustments in that they stay too close to the original anchor. To model this we assume that people start with the anchor L (the amount to lose if the hazard occurs) and adjust this amount downward aiming at the target pL. We further assume a consistent proportional underadjustment error of β. If the total adjustment ought, according to an EV model, to be (L-pL) then the error will be $\beta(L-pL)$. In other words, the bid will be given by

$$B = pL + \beta(L-pL).$$

Dividing both sides by EV = pL yields

$$B/EV = 1 + \beta(1-p)/p.$$

Fitting this model to the mean bids for those people not editing gives the following estimates:

$$B/EV = \underset{(13.3)}{1.1} + \underset{(11.5)}{.023}(1-p)/p$$

$$DF = 4 \qquad R^2 = .97$$

The intercept of 1.1 is not significantly different from the predicted value of 1 (t = 1.2). The data and the model fit are plotted in Figure 3.4.

Surprisingly, the model fit estimates the underadjustment error as only being between 2% and 3%. How can such a small error distort responses so much for low probabilities? For low probabilities the absolute size of the adjustment is large because the loss serves as the anchor and the distance between that anchor (L) and the target (pL) will be greatest for low probabilities. As the size of the adjustment required becomes bigger so too does the size of the adjustment error B(L-pL). This absolute adjustment error will be especially large when considered relative to pL, the EV target, which will be small for low probabilities. Hence, overbidding relative to EV will be especially pronounced, according to the anchoring and adjustment model, for low probabilities.

From subsequent laboratory experiments it appears that the anchoring and adjustment process may also apply to those who edit the risk, that is people who edit actually select O as their anchor and then adjust upward to reflect the probability and magnitude of the consequence. In such cases, the target is close to the anchors so adjustment errors may be relatively very small.

If bimodality were only found in the laboratory it would be of theoretical but not practical interest. However, we found the same bimodality of risk judgments in our field survey of residents living near the Operating Industries, Inc. (OII) landfill in California. Figure 4.3 of Chapter 4 shows the frequency distribution for resident's subjective health risk beliefs (again on a log scale). The bimodality is apparent: a large proportion of people essentially dismiss the risk by giving an estimate approximately equal to the probability of death from saccharin consumption while others give a much higher estimate approximating the risk of death from smoking a pack of cigarettes per day. We believe those equating the landfill risk with saccharin consumption are editing and then adjusting upward from a zero anchor while those equating it with lung cancer from smoking are adjusting downward from a cancer death anchor. Epidemiological studies and on-site monitoring of hazardous chemicals fail to reveal any significant risks at OII. Thus, those neighborhood residents in the low risk mode parallel those participants in our lab studies who essentially dismissed or edited the risk while those residents in the high risk mode parallel those lab participants who had an exaggerated concern for the low probability risk. We therefore think that the cognitive processes of editing and anchoring and adjustment are operating in the field sites.

When we have presented this research at professional meetings, other researchers have always come forward to tell us of similar bimodal results in the area of health risks which heretofore had been puzzling. We are convinced that it is a ubiquitous phenomenon.

An especially noteworthy aspect of the OII field study was the demonstration that the higher the proportion of residents in a neighborhood who were in the high risk mode the greater the decrease in property values. Thus, these subjective risk estimates had a real impact on economic behavior. Note also that bimodality has important implications for community conflict. It suggests there will be two distinct community perspectives towards low probability risks. One group complains that the risk is negligible and that all the fuss will only lower property values while the other group cannot understand why the former group only worries about money and is not concerned about the deadly risk confronting them all. Such conflicts are typical of low-risk Superfund sites. In summary, when deciding how concerned to be about a low probability, high consequence risk, the first decision is whether to give the risk any attention at all. If not, the risk is dismissed or edited and the person has virtually no concern for the risk. However, if the risk is not edited, in deciding how concerned to be, the person starts with his or her concern with the consequence if it were to occur for certain. Then the person adjusts that concern downwards to reflect the relative likelihood of the risk. This adjustment will almost always be an underadjustment. Given that the consequence will usually be very extreme (e.g., death from cancer), even a, slight underadjustment will leave the person with a relatively high level of concern. The operation of both editing and anchoring and adjustment processes produces a bimodal distribution of concern. As a shorthand in the following, we will refer to those people who dismiss a low probability risk as being in the “edit mode” and to those who are overly concerned as being in the “concern mode."

2.3 Two kinds of Low Probability Risks.

Given the bimodality and the difficulty that people have dealing with numerical risk estimates, especially for very low probabilities, we are doubtful that any attempts to communicate precise quantitative risk estimates will be successful. About the most that can be hoped for, in our opinion, is to get people to have levels of concern in the more appropriate mode, realizing that neither mode may be accurate. That is, if information about a new risk is being presented, the communication effort should be directed at getting people in either the edit mode or the concern mode depending on whether or not it is important to get people to take action. For existing risks for which there are some people in each mode, communication might be directed at moving people from one mode to the other.

The possibility of moving from one mode to another suggests that there is an important distinction between risk types. The first type is those risks for which it is appropriate to have communications attempting to move people from the concern mode to the edit mode. Examples might include Superfund sites where there is a great community concern but little or no scientific evidence of a risk. The second type is those risks for which it is appropriate to have communications attempting to move people from the edit mode to the concern mode. Examples might include efforts to get people to test their homes for radon gas accumulations. The communication strategies will likely be very different for the two kinds of low risk problems.

2.4 Determinants of Risk Mode.

Given the importance of the two risk modes, it is necessary to consider the factors which influence the mode in which a person’s risk judgments fall. Knowledge of those factors will automatically suggest strategies for moving people from one mode to another. In this chapter we try to answer why some low probability risks are generally ignored or edited while others generate great concern.

A dramatic example is presented by homeowners in New Jersey. Even though their homes are at high risk of having significant radon gas levels, it is extremely difficult to get homeowners in that area to test for radon, let alone make any building modifications to reduce radon gas accumulation. However, some of those same homeowners are greatly concerned about barrels of low-level radioactive wastes from the former manufacturer of radium watch dials in a nearby factory. Ironically, some of these homeowners would probably be less at risk if they were to live in a tent in the storage yard housing the barrels of low-level radioactive wastes than to sleep regularly in their basements.

In this subsection we consider three classes of factors that have been shown in our research and that of other’s to influence a person’s risk mode: perceptual cues, consequences, and experience.

2.4.1. Perceptual Cues.

Risk judgments are often partly based on perceptual cues. The more people are reminded of a risk, the more likely they are to be in the concern risk judgment mode. For example, residents near a landfill may form risk judgments based on such cues as foul odors from the site, heavy traffic to and from the site, and the presence of chain link fences. Combined with media attention indicating a possible offsite problem with cancer, these perceptual cues may move many residents to have a high concern about the risk of living near the site. In contrast, explosive concentrations of methane gas escaping from a landfill do not possess any perceptual cues to warn people of the dangers. The gas is odorless, colorless and tasteless. People depend totally on some form of official warning to take action. Fortunately, for low probability risks, many characteristics can be manipulated so that the public may evaluate such risks more accurately. In this context it is interesting to note that natural gas companies add an artificial odor to natural gas so that residents can easily detect the risk if natural gas is leaking in their homes.

Perceptual cues may be classified according to the five senses of sight, smell, taste, hearing and touch. Of these, visible cues are probably the most easily controlled. Often protective measures inadvertantly alarm the public to an excessive degree. For example, instead of institutional looking chain link fences, hedges could be planted to make a low-hazard area more aesthetic.

Other controllable perceptual cues include smell and taste. People who are bothered by odors emanating from a site such as a landfill or chemical facility are more apt to perceive greater risks. For example, in Chapter 4 we observe that residents near the OII landfill were made fearful by such odors. Another specific illustration is iron contamination in the Eagle River from the Eagle Mine which resulted in discoloration and poor taste of drinking water in Vail, Colorado. Local residents were greatly alarmed and feared that their water supply had been poisoned even though the iron in the water was itself non-toxic.

The media also provide many perceptual cues through written or spoken statements and still or moving photographic images. Our research at OII indicated that frequent exposure to media reports about OII increased the likelihood that someone would be in the concern mode.

In summary, the presence of strong perceptual cues moves risk judgments to the concern mode while the absence of perceptual cues allows risks to be edited more easily. Clearly, a key component of any risk communication strategy would be to add or reduce perceptual cues, depending on whether it was appropriate, respectively, to increase or decrease concern about a particular risk.

2.4.2. Consequences.

The nature of the consequences, or more accurately, a person's beliefs about and characterization of those consequences, are obviously important in determining the risk mode. Clearly, the more serious and dramatic the consequences of a risk, the higher will be the anchor in the anchoring and adjustment process so the final level of concern will be higher. Slovic, Fischhoff, Lichtenstein (1980) have identified a number of important characteristics of risks that cause people to under- or overestimate risks, or, in our terms, to be in the edit mode or the concern mode. Many of these characteristics pertain to the magnitude and imagibility of the consequence. Dreaded risks that are believed to have the potential of killing many people at one time in a dramatic event will usually have risk estimates in the concern mode.

One very important characteristic of a risk's consequence that will strongly affect attempts at risk communication is whether people believe the change in risk is a loss or a gain relative to the status quo. The differentiation of loss and gain effects is based on Kahneman and Tversky's prospect theory. As a rough rule of thumb, a loss relative to the status quo will have about three to ten times the psychological impact as an equivalent gain. Going from thinking you are safe to believing you are unsafe makes people a lot more unhappy than going from unsafe to safe makes people happy. In the context of hazardous wastes, for example, informing residents about an old chemical waste dump in their neighborhood about which they had no awareness will create a great deal of unhappiness because it is a loss relative to the status quo. Conversely, telling people who have worried about a known waste site for many years that it is in fact safe, even if they were to believe it, would not increase their happiness much because it is a gain relative to the status quo. This means that informing people about new risks and hazards must be done very carefully and that informing people about old risks is not likely to have much impact.

We believe this distinction between gains and losses partly explains the usual disparity between willingness-to-pay (WTP) and willingness-to-accept (WTA) measures of concern about risks. Fisher, McClelland, and Schulze (1986) review a number of studies comparing WTP and WTA responses. Of particular interest is our OII study (summarized in Chapter 4) which found that for closing the site if it were open (again) people were willing to pay (WTP) only about a tenth as much as they demanded (WTA) to allow the site to reopen (a loss) if it were closed.

It is important to note that there are at least two important ways in which a risk could be viewed as a loss relative to the status quo. First, people may learn of the probability or possibility of a risk that they previously did not believe they were exposed to. Second, people may learn that the consequences of a risk are more severe than they realized. Either or both of these will create an exaggerated concern about the risk.

2.4.3. Experience.

The amount and nature of prior experience with the risk is an important determinant of whether a person will be in the edit or concern risk mode. Research by Slovic et al. (1980) and others has shown that risks that are familiar, well-known to science, and with which we have had lots of benign experience are more likely to be edited while risks that are unfamiliar, not well understood by science, and with which we don’t think we have had benign experiences are more likely to generate high levels of concern. For example, almost everyone, even those who have had automobile accidents, has had numerous benign driving experiences. Hence, the risks of automobile driving tend to be underestimated or edited, resulting in the underuse of seat belts. On the other hand, people who have not flown much, if any, have not had benign experience with airplanes and therefore tend to overestimate the dangers of air travel.

In our laboratory studies we have been able to take a close look at the effect of experience on risk mode. In several studies we have looked at the concern (as reflected in bids for insurance) about a low probability risk when people are exposed to many rounds of the risk (either 50 or 150 rounds). Figure 3.10 in Chapter 3 shows a 50 round experiment where the concern mode steadily decreases with benign experience until the risk occurs on the 33rd round. At that point, there is a sharp drop in the proportion in the concern mode reflecting the gambler’s fallacy that a low probability event is less likely to occur on the next round because it occurred on the previous round. But then for succeeding rounds the number of people in the concern mode grows as fewer and fewer people feel comfortable in editing the risk. Note that there is not a corresponding sharp drop in the average bids for those bidding after the first risk event at round 33. This suggests that the effect of experiential variables such as benign experience and the gambler’s fallacy are more important in determining whether someone edits or not than in determining the actual level of their concern.

The three factors that influence whether a person is in the edit or concern mode can explain the ironic, paradoxical response of those homeowners in New Jersey. They have all had lots of benign experience living in their homes; it is hard for them to imagine that they will be harmed in their own homes and that it will be their home that does the harm. Deaths from the radon gas risk will be undramatic and difficult to attribute to radon. Also, there are no perceptual cues because radon gas is invisible, odorless, and tasteless. Hence, the radon gas risk is underestimated, probably seriously underestimated because so many people edit the risk. On the other hand, residents are likely to believe, falsely, that the radioactive barrels might explode and wipe out the neighborhood in a dramatic event; the well-known radioactivity sign provides a perceptual cue; residents are also unaware of their benign experience living with this risk; and they don't think that the radioactivity is well understood by scientists. Hence, it is a risk that will be overestimated.

Summary. Our basic model is that perceptual cues to the risk, the risk consequences, and experience with the risk determine whether people dismiss or edit the risk. If they edit the risk, they will show little or no concern for the risk. If they don't edit, then an anchoring and adjustment model describes how people arrive at their level of concern for the risk. Anchors will generally be the consequence of loss, and adjustments downward from this anchor to reflect the low probability will almost always be underadjustments resulting in an exaggerated level of concern. The operation of the editing and anchoring and adjustment cognitive processes results in a bimodal distribution of concern levels for low probability, high consequence risks. The best that may be hoped for in risk communication is to move people to either the edit or concern mode, depending on which mode is closer to the appropriate level of concern for that risk.

2.5 Changing Responses to Low Probabilities.

The literature review and especially our own research suggests that the primary problems in risk communication involve low probability, high consequence situations, which are typical of many Superfund sites. In this subsection we describe some preliminary research which suggests some strategies for changing responses to low probabilities. The studies reported here are indeed preliminary and do not yet warrant their own separate technical reports or chapters in this report.

2.5.1 Changing Low to High Probabilities.

Our research and that of others suggests that although people have difficulty understanding the implications of low probabilities for appropriate behavior, they have little or no difficulty with moderate to high probabilities. One way to change an apparent low probability of risk on annual basis to a moderate probability is to change the focus from an annual basis to a longer period such as a lifetime basis. For example, for an annual risk of .01, the risk for a 70-year lifetime would be $1-(1-.01)^{70} = .51$. Our research presented in Chapter 3 suggests that while people would have difficulty understanding the .01 annual risk they might be better prepared to understand the .51 lifetime risk. Slovic, Fischoff, and Lichtenstein (1978) report suggestive evidence that such a lifetime focus increases willingness to use automobile seat belts.

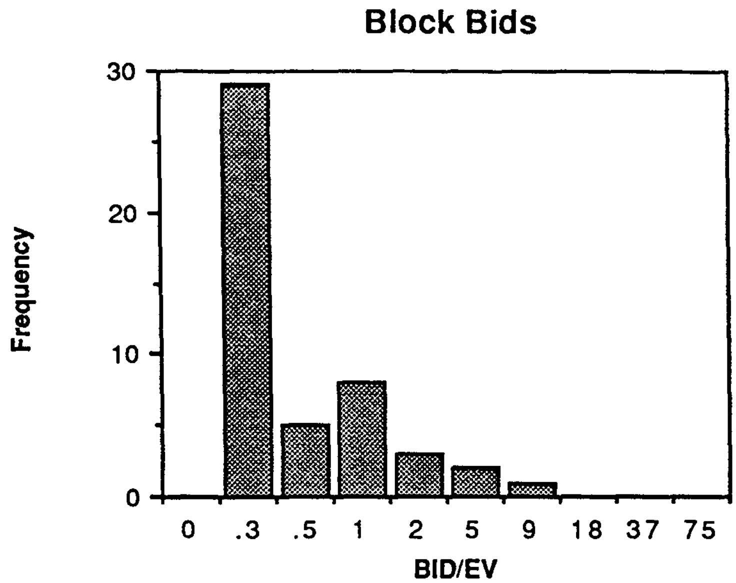
To test this idea more rigorously, we have modified the experimental procedure used in Chapter 3. Participants were told that the probability of loss on any given round was .01 but that the probability of at least one loss during 25 rounds was approximately .22. Participants were then given the opportunity to bid for an insurance policy that would protect them against a loss for the entire 25 rounds. After the first block of

25 rounds (in which no loss was experienced), they had the opportunity to bid for a second insurance policy that would protect them against a loss for the second block of 25 rounds. For comparison, the round-by-round bids of a different group of participants were summed across each block of 25 trials. In the first block of 25 rounds, for those bidding round-by-round mean BID/EV = 6.5, consistent with our earlier studies; however, for those bidding for a block of 25 rounds mean BID/EV was only 0.86. For the second block of 25 rounds the respective means were 8.5 and 0.63. Thus, even though the objective risk situation was identical for both groups of participants, changing the focus from round-to-round to blocks of 25 rounds dramatically changed behavior, as expected from our model, in a direction towards EV. Somewhat surprisingly, many participants were substantially below EV indicating that the proportion editing (or more likely anchoring on zero) increased with the block-of-25 focus. Figures 2.1 and 2.2 display the comparison of the frequency distributions of BID/EV. The block-of-25 focus shifts the entire distribution downward, note especially that extremely high bids (relative to EV) are eliminated by the block-of-25 focus.

Our results , consistent with Slovic, et al. (1978) suggest that an effective strategy in communicating low probability risks is to use a long enough focus so that the probability will be at an understandable level. For annual risks on the order of 10^{-2} and maybe even 10^{-3}, a lifetime focus might be useful or other time frames such as the “length of your mortgage.” Unfortunately, this strategy will not work for risks of 10^{-4} or lower because even the lifetime risk is not in a range of probabilities that most people will be able to understand. One possibility might be also to change the focus from the individual to the family or community. For example, for

FIGURE 2.1

FREQUENCY DISTRIBUTIONS OF BID/EV FOR ROUNDS 1-25



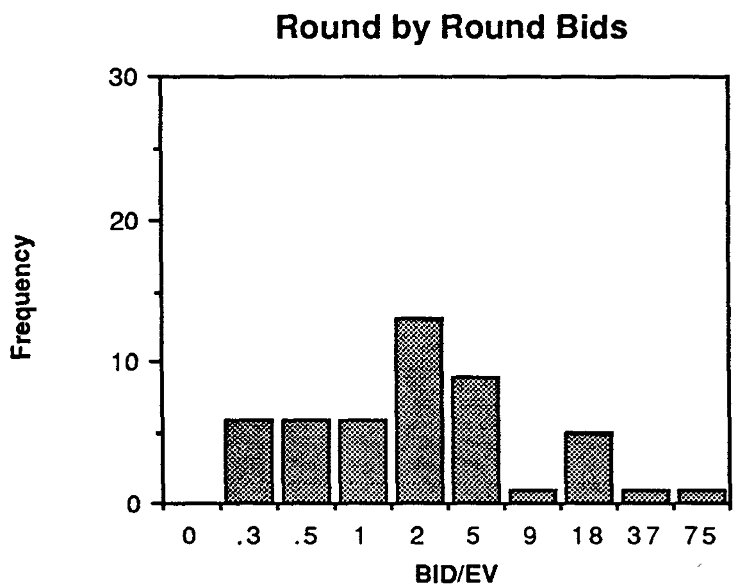


FIGURE 2.2

FREQUENCY DISTRIBUTIONS OF BID/EV FOR
ROUNDS 26-50

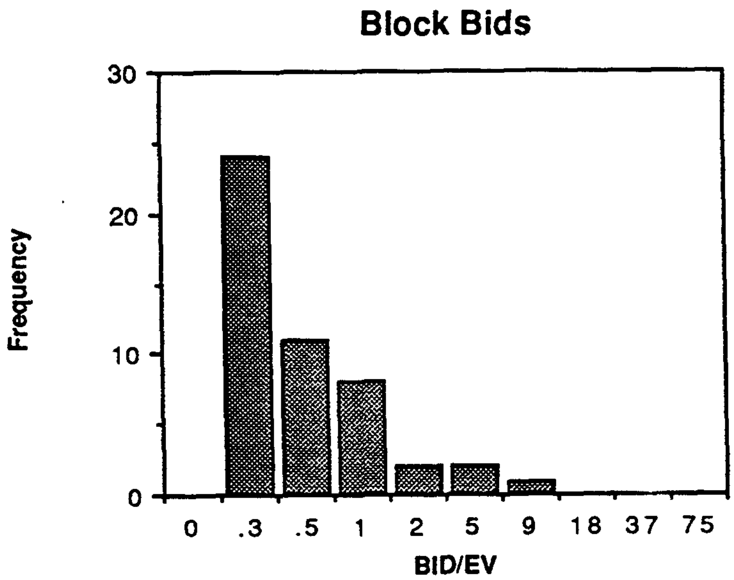
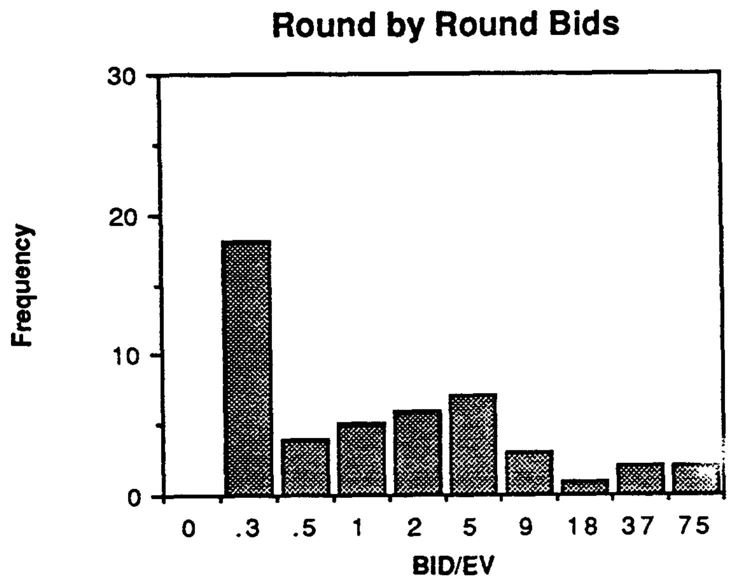


FIGURE 2.2
(CONTINUED)

FREQUENCY DISTRIBUTIONS OF BID/EV FOR
ROUNDS 26-60



an annual risk of 10^{-4}, the probability that over a lifetime at least one member of a family of four will suffer the risk is much more understandable than the annual risk. Similarly, in a neighborhood of 100 people, the probability that at least one person would suffer the risk over a lifetime would be approximately .5. The neighborhood should be able to understand the .5 probability and respond appropriately. However, this is a speculative conclusion and should not be implemented until more research on the effects of changing focus have been conducted.

Section 2.4 reviewed the findings on determinants of risk mode--those factors which influence whether people dismiss particular low probability risks or develop a high level of concern about them. It is natural to look to those factors for potential mechanisms for changing responses to low probability, high consequence events. We consider each of the three types of determinants in turn.

2.5.2 Perceptual Cues.

The perceptual cues about a hazard such as a Superfund site are just as important a part of the risk communication process as any of the formal intended messages sent to neighborhood residents by public officials and the media. Thus, it is critically important that as much as possible the perceptual cues are consistent with the actual level of risk. Too often, at sites with very low risk levels, perceptual cues such as high chain link fences, danger signs, and odors send a message inconsistent with the actual risk. On the other hand, if people are not taking sufficient precautions near a Superfund site, it would be advisable to increase the perceptual reminders about the hazard. It would be interesting to test these notions about changing perceptual cues in the laboratory but we have not yet had the opportunity to do so.

Cutter (1987) has analyzed the case histories of the evacuation response for several large accidental airborne toxic releases such as the derailment and chlorine release at Mississauga, Ontario in 1979. She too points to the important role of perceptual cues in citizen judgments of risk. Those residents who have perceptual cues about an accident (see smoke, hear an explosion, smell unusual odors, see and hear emergency equipment, etc.) are easy to evacuate; in fact, those residents often evacuate before asked to do so. On the other hand, those residents who live close enough to be in danger but far enough away so that perceptual cues are weak or nonexistent are often very reluctant to evacuate even when ordered to do so by uniformed policy officers. Cutter worries that sometimes residents over-respond to perceptual cues that might cause panic and suggests that efforts be made to bring perceptual cues into line with the actual risk as much as possible.

2.5.3 Consequences.

How citizens view the potential consequences of a risk event is an important determinant of the magnitude of their response. As noted above, of particular importance is whether a change in risk is viewed as a gain or a loss relative to the status quo. Sometimes it is possible to present risk information in a form that suggests a reference frame. For example, McNeil, Pauker, Sox, and Tversky (1982) asked both physicians and patients to choose between two treatments: an operation involving some risk of perioperative death and a less effective treatment not involving an operation. For half the people the probability was expressed in terms of dying (e.g., 5 out of 100 will die); for the other half the probability was expressed in terms of living (e.g., 95 out of 100 will live). Even though the risk was the same, both physicians <u>and</u> patients more often selected the risky operation when its risk was expressed in terms of the number living instead of the number dying. In a similar study involving the risks of oral contraceptives, Halpern, Blackman, and Salzman (1986) presented women with the same risk information expressed in these two forms: In the age group 15 to 34 it has been estimated that, in general, about 99991.7 out of every 100,000 women who use oral contraceptives will not die of circulatory disorders each year; whereas, for nonusers the rate is about 99998 out of 100,000 each year who will not die of circulatory disorders. In the age group 15 to 34 it has been estimated that, in general, about 8.3 out of every 100,000 women who use oral contraceptives die per year; whereas, for nonusers the rate is about 2 out of every 100,000 per year. Women who received the first communication judged that oral contraceptives were less risky than did the women who received the second communication. These studies suggest, therefore, that expressing risks at a Superfund site in terms of the probability that there will not be an accident or the probability that no one will be hurt by the site is likely to generate lower levels of concern than expressing those risks in terms of the probability that there will be be accident or that someone will be hurt by the site.

Some preliminary laboratory experiments we have conducted suggest that reference frames (viewing something as a loss or a gain) may change with time. We made a slight modification of our standard procedure (see Chapter 3, for a complete description of the standard procedure) so that instead of buying insurance for a low probability, high consequence risk, participants were given insurance with the opportunity to sell it back to the experimenter. That is, the procedure was changed from willingness-to-pay (WTP) to willingness-to-accept (WTA) or compensation demanded (CD). On early rounds some participants demand very high prices for their insurance but in later rounds suddenly switch to much lower prices more consistent with the expected value of the insurance. Our hypothesis is that they initially adopt the frame that they have the insurance policy, a valuable commodity, and that to give it up will be a loss. Hence, they demand a high amount for their insurance. But we buy back the insurance only from the four (out of eight) lowest bidders so those demanding high amounts keep their insurance. They then come to view the money that others are getting for selling their insurance as money that they themselves could have had; hence, the money not obtained becomes viewed as a loss. From this new frame or perspective, participants then quickly lower their offering prices so as not to miss out on the money from the experimenters.

It is possible that reframing of gains and losses similar to our laboratory studies might take place in the context of a Superfund site. For example, consider people living near a site who are so concerned about possible water contamination that they use bottled water exclusively. As they see that no problems are occurring to their neighbors who are not engaging in such averting behaviors, they may come to see their own extra efforts as a loss. Similarly, if a lot of money were spent on a detection and warning equipment by a municipality and a problem never materialized, that money and any future expenditures would likely be viewed as a loss. However, this remains an interesting speculation that ought to be tested empirically.

2.5.4 Experience.

With respect to exaggerated concern about a risk and community conflict about the level of a risk, the aphorism that "time heals all wounds” does appear to apply. With time, as long as the risk remains a potential risk and nothing happens, concern about the risk will generally decrease. The operating mechanism is of course not time; instead, the reduction in risk judgments is probably due to increasing familiarity. If so, this suggests that an amelioration strategy is to increase community familiarity with the Superfund site. Thus, in those cases where it was safe to do so, someone ought to conduct tours of a site so that people can become more familiar with the situation. In such cases, people's worst imaginations are almost always worse than the actual fact so becoming more familiar with the risk almost always reduces their level of concern.

CHAPTER 3: AN EXPERIMENTAL STUDY OF FACTORS AFFECTING RISK BELIEFS

3.1 Introduction

Psychologists have documented many systematic deviations in behavior from that predicted by the expected utility model. Much of this evidence has been generated in experiments in which subjects have been asked what their behavior would be in response to hypothetical situations (see, for example, Lichtenstein and Slovic, 1971; Slovic et al., 1977; Kahneman and Tversky, 1979; Tversky and Kahneman, 1974, 1981; Abelson and Levi, 1985). Based on these experiments, psychologists have argued that errors in decision making under uncertainty arise from the improper application of intuition or simplifying rules of thumb (heuristics), from the improper consideration of factors irrelevant to the decision (framing or context effects), and from errors in reasoning about probabilities. Such errors may play a dominant role under some circumstances such as those found at Superfund sites.

Economists have also conducted laboratory experiments exploring behavior under uncertainty. Results of these experiments, while in part confirming deviations from the expected utility model (e.g., Grether and Plott, 1979), suggest that when individuals make repeated choices in a market environment the expected utility model is "not universally misleading" (Plott and Sunder, 1982, p. 692). Economic experiments generally use actual cash payments, induce values (control the value to the subject of the commodity used in the experiment so it is known with certainty to the experimenter (Smith, 1976), and employ many repeated trials to allow individuals to practice and become familiar with the market institution (e.g., Coppinger, Smith and Titus, 1980; Smith, Williams, Bratton and Vannoni, 1982; and Coursey, Hovis and Schulze, 1986).

One principal focus of experimental economics has been an examination of the efficiency and Pareto optimality properties of market institutions. Since Pareto optimality by definition is an idealized rational outcome, experimental economists have been concerned with finding institutions which tend to produce rational behavior. This focus contrasts substantially with the objective of many experiments, conducted by psychologists, that have as their objective the detection of situations where deviations from rational behavior will occur.

This chapter presents results of two experiments that attempt to integrate these separate lines of research conducted by economists and psychologists to understand behavior at low probabilities. To this end, our experimental design and our analysis follow procedures and employ concepts drawn from both cognitive psychology and experimental economics.

The aim is to collect a body of evidence which might help in interpreting the empirical study of a Superfund site presented in Chapter 4. This study suggests that for low probability, high loss events, large deviations from rational behavior are likely to occur. For example, past studies of flood and earthquake insurance (Kunreuther, et al., 1978) and of the value of avoiding exposure to hazardous substances (Burness et al., 1978 and Smith and Desvouges, 1986) all suggest deviations from rationality. Such studies document a difficult and as yet unresolved policy dilemma. In some cases, such as hazardous wastes, many individuals seem to place inexplicably large values on avoiding risks. Yet in other cases, such as floods or earthquakes, many individuals refuse to insure against objectively similar or even greater risks. While it is difficult or impossible to replicate the high loss nature of such events in the laboratory, it is possible to explore a range of risk to see if behavior at relatively lower probabilities is in some way different from behavior at relatively higher probabilities.

Finally, considerable controversy has surrounded the use of hypothetical as opposed to actual responses from individuals. Thus, the experiments were also designed to collect both hypothetical and actual data involving cash purchases of insurance. Hypothetical values were obtained both before and after individuals had actual repeated market-like experience so that the effect of experience on the accuracy of hypothetical responses could be assessed.

In interpreting results such as those reported here, there is always some question about how well responses to laboratory risks generalize to real risks posed by a Superfund site. Although the precise responses might not generalize there is good reason to expect that the cognitive processes underlying the responses will generalize. That is, while it may not be possible to predict from laboratory studies the proportion of people who edit or the amount by which people underadjust from an anchor or the degree to which they are susceptible to the gambler’s fallacy, it is almost surely true that the cognitive processes of editing, anchoring and adjustment, and the gambler’s fallacy will be evident in people’s responses to a Superfund site. It is unreasonable to presume that people in emotionally-charged situations attempting to deal with real risks will suddenly have access to cognitive processes to aid them in their decisions that they do not have access to in the relative calmness of the laboratory. It is our view that, laboratory. It is our view that, if anything, the decision making problems identified in the studies reported here are likely to be exacerbated when people confront risks outside the laboratory.

3.2. Experimental Design

3.2.1 Theoretical Issues

This section develops the theoretical basis for the detailed experimental design presented below. In contrasting expected utility theory (EUT) with models from cognitive psychology, we draw strongly on the formalized theoretical structure developed by Kahneman and Tversky (1979) which they term prospect theory (PT). PT has been evolving over the last decade and we apply the label broadly to include several extensions of the model.

In general, our experiments were conducted as follows: each subject is given the opportunity to make a bid of B dollars for insurance against a possible loss of L dollars that occurs if a red chip is drawn. The probability of drawing a red chip is given as p. If a white chip with a stated probability of 1-p is drawn, each subject is rewarded with a gain of G dollars. The gain is included in part to finance successive trials. If a subject has an initial wealth of $Y°$ dollars and utility is a function U(Y) of wealth Y, then, according to EUT the expected utility of the situation described above without purchase of insurance is

(3.2.1) $pU(Y°-L)+(1-p)U(Y°+G)$

and the expected utility with purchase of insurance is

(3.2.2) $pU(Y°-B)+(1-p)U(Y°+G-B)$.

The most that an individual should pay for insurance can be obtained by setting (3.2.2) equal to (3.2.1) and solving for the bid, B. The notion here is that individuals will only be willing to increase the bid to the point that the expected utility with insurance falls to the level of of expected utility without insurance. Since the loss and gain used in the first experiment ($4 and $1, respectively) are small relative to wealth, EUT would imply that it is reasonable to suppose that changes in wealth are constrained to an approximately linear segment of the utility function. Thus, a linear approximate utility function

(3.2.3) $U(Y)=U(Y^o)+U'(Y^o)\cdot\Delta Y$ where $\Delta Y=Y-Y^o$

may be substituted into (3.2.1) and (3.2.2) without loss of generality. If (3.2.1) and (3.2.2) are then set equal, the bid for insurance solves as

(3.2.4) $B=p\cdot L.$

Thus, the bid is equal to the expected value of the loss (EV). Since, as noted above, Vickrey auctions have been shown to be strongly demand revealing, we would expect bids to be equal to EV or at least normally distributed around EV for a large range of probabilities if EUT is a good predictor of behavior.

While maintaining the linear weighting of EUT, prospect theory makes two modifications. First, the utility function is replaced with a rather different <u>value function</u>. Second, the probabilities are replaced by a <u>weighting function</u> which depends on the probabilities.

PT postulates that individuals are assumed to care only about relative changes from their current wealth position and to dislike a loss in wealth much more than they enjoy an equivalent gain. Thus, according to PT the value function is not an argument of wealth, but rather of changes in wealth, AY. Further, the value function $\nu(\Delta Y)$ has the properties that $\nu(o)=0$, the left hand derivative $\nu'(o)^-$ exceeds the right hand derivative $\nu'(o)^+$ at the origin, and that both derivatives are positive, so $\nu'(o)^->\nu'(o)^+>0$. As we show below, the value function likely plays no role in the structure of our experiment, but it has been introduced by cognitive psychologists because many individuals seem to make errors in judgment because they reason in relative rather than absolute terms and show intense aversion to perceived losses.

The weighting function $\pi(p)$ of PT overweights small probabilities ($\pi(p)>p$), underweights large probabilities ($\pi(p)<p$) and shows subcertainty ($\pi(p)+\pi(1-P)<1$). The subcertainty feature implies that when a certain outcome is compared to an uncertain prospect, the prospect will be underweighted relative to the certain outcome. This attribute of the model adjusts for the observation drawn from psychology experiments that individuals seem to be biased towards certainty. Similar probability weighting functions have been proposed by Handa (1977) and Karmarkar (1979).

Given PT as described above, the value of the prospect posed by the experimental situation without insurance would be given by

(3.2.5) $\pi(p)\nu(-L)+\pi(1-p)\nu(G)$

and the value of the situation with insurance would be given by

(3.2.6) $\nu(-B)+\pi(p)\nu(o)+\pi(1-p)\nu(G)$.

Note that (3.2.6) is not written as $\pi(p)\nu(-B)+\pi(1-p)\nu(G-B)$. This is because subjects must first pay for insurance, a certain loss which is valued as $\nu(-B)$ and implicitly weighted with unity. After this adjustment, subjects face a modified prospect of $\pi(p)\nu(o)+\pi(1-p)\nu(G)$ which is underweighted since $\pi(p)+\pi(1-p)<1$, reflecting a bias against uncertainty central to PT.

To obtain the bid for insurance, the two expressions (3.2.5) and (3.2.6) are set equal. This algebraic manipulation is specifically legitimized by cognitive psychology in the following way. The model presented here can be interpreted as a mental representation that

that describes how individuals decide how much to bid for insurance. Thus, subjects in the experiment will note that the gain of G dollars will occur with or without purchase of insurance. This implies that $\pi(1-p)\nu(G)$ may be cancelled from (3.2.5) and (3.2.6), that is, the gain can be ignored in the decision process. If an individual has insurance, a red draw causes no loss, so the term $\pi(p)\nu(o)$ may be dropped from (3.2.6) since $\nu(o)=0$. 'his leaves a comparison of the certain loss associated with purchasing insurance which is valued as $\nu(-B)$ with the uncertain loss associated with drawing a red chip which is valued as $\pi(p)\nu(-L)$. Thus, we arrive at

(3.2.7) $\nu(-B)=\pi(p)\nu(-L)$.

Since the value functions on both sides of (3.2.7) evaluate small decreases in income, -B and -L, respectively, a linear approximation of the value function is appropriate so, for decreases in income ($\Delta Y<0$) we have

(3.2.8) $\nu(\Delta Y)=\nu(o)+\nu'(o)^{-}\cdot\Delta Y=\nu'(o)^{-}\cdot\Delta Y$

since $\nu(o)=0$. Substituting (2.8) into (2.7) yields

(3.2.9) $B = \pi(p)L$

and therefore the bid is equal to the weighting function times the loss. Thus, bids for insurance against a small loss will, according to PT, involve the weighting function but not the value function. Individuals can be thought of as recognizing that they must choose between two small losses: a sure one of B dollars and an unsure one of L dollars. We will discuss a possible mental process for arriving at this bid later.

In analyzing the data from the experiments we can evaluate the predictions of of relative to those of EUT by dividing the actual bids obtained in the experiment by EV which is a known constant, pL, for any stated probability, P, and loss, L. If PT is taken as the basis of analysis, dividing (3.2.9) by EV yields

(3.2.10) $B/EV=\pi(p)/p$.

Given the assumptions on the weighting function (relative overweighting of low probabilities) B/EV should be greater than unity for small probabilities, and B/EV should be less than unity for larger probabilities. Thus, our experimental design focuses on the values of B/EV over alternative probability levels. If the frequency distribution of individual values of B/EV at all probability levels is normally distributed around unity, then bids should closely correspond to EV and EUT would be supported by the data. However, if the frequency distribution of individual values of B/EV is not normally distributed around unity, some alternative, such as PT, is likely to be the more appropriate theoretical structure.

3.2.2 The Structure of the Experiment

Each experimental session employed eight student volunteers recruited from undergraduate economics classes at the University of Colorado. Five experimental sessions (total of 40 participants) provide data at probabilities of .01, .1, .2 and .4 while three experimental sessions (total of 24 participants) provide data at probabilities of .6 and .9. No student participated in more than one session. Subjects received a $5 guaranteed payment for participating. In addition, they were given a $10 stake at the beginning of the five lower probability experimental sessions and a stake of $65 at the start of the three higher probability experimental sessions. They were allowed to keep any of the stake remaining and any gains at the end of the experiment. Subjects were assured that even if they lost all their stake, they would still receive the $5 payment.

Overview. In the course of the five lower probability experimental sessions, each participant made a total of 51 bids to purchase insurance in the following risky situation which was fully described to the participants: A chip is to be drawn from a bag containing R red chips and W = 100-R white chips. If a white chip is drawn, each participant receives $1. If a red chip is drawn, those having insurance lose nothing but those without insurance lose $4. Before being placed in the bag, the stacks of chips were displayed on a table in front of the participants so they would have a more concrete representation of the specific probability levels. The four values of R used in each session were 1, 10, 20, and 40 corresponding to p = .01, .1, .2, and .4, respectively. The particular value of R being used was always made explicit before each bid. The total of 51 bids consisted of two basic types: hypothetical bids (7) and Vickrey auction bids (44). The method used for obtaining each bid type is described separately below and then the sequence of the bid types is described. In the three higher probability sessions an identical situation was employed where, however, R was equal to 60 and 90 corresponding to probabilities of .6 and .9, respectively. Three hypothetical and 20 actual bids were collected from each subject.

Hypothetical Bids. Two types of hypothetical bids were collected: inexperienced and experienced. For the inexperienced hypothetical bids, the risky situation was described to subjects as hypothetical and they were asked how much they would hypothetically pay for an “insurance policy”, which would offer full protection against the $4 loss associated with the draw of a red chip. Subjects wrote their bids on paper. These inexperienced hypothetical bids were meant to be comparable to the types of responses obtained in many psychology experiments (for example, Slovic & Liechtenstein, 1968).

To obtain the experienced hypothetical bids, subjects were asked the same hypothetical question <u>after</u> they had had experience with the Vickrey auction and with the drawing of chips for other probability levels. Subjects entered their bids on computer terminals in the same manner as described below for the Vickrey auction.

<u>Vickrey Auction Bids</u>. A Vickrey (1961) auction determined who received insurance on each round. Subjects read written instructions, heard an oral explanation of the auction procedure, and were given an opportunity to ask questions. After the appropriate number of chips were displayed and placed in the bag, the eight subjects in each session entered bids on a computer terminal for one of four insurance policies sold in each round. The terminal also displayed the current composition of the chip bag. The computer accepted bids between, inclusively, 0 and the subject’s current balance in units of one cent. After everyone had entered a bid, the computer rank ordered the bids from highest to lowest and displayed the "reigning price”--the fifth highest bid for insurance--on each subject's terminal screen. Only the four subjects with bids above the reigning price received insurance. In the case of ties for the fourth highest bid, remaining insurance policies were randomly allocated among those with tied bids. Those receiving insurance were only required to pay the reigning price. This represents the key feature of the Vickrey auction and is intended to eliminate the incentives for strategic behavior that are present in auctions in which individuals must pay exactly what they bid. After each auction, the computer displayed the original balance, the reigning price, whether or not insurance had been received, adjustments to the balance, if any, and the new balance. Other than the reigning price,

subjects received no information about the bids of other subjects. Terminals were arranged so that no subject could see the terminal of any other subject and subjects were not allowed to talk with each other. At the beginning of the experiment subjects participated in four practice bidding rounds which did not affect their balances in order to familiarize them with the procedures used in the Vickrey auction.

Great care was taken to avoid the use of any judgmental words in the written and oral instructions. This is in contrast to some previous experiments using the Vickrey auction which have used “winners” to designate those who have received insurance. The use of such words might artificially increase the subjective value of holding insurance above its value as protection against the loss associated with the draw of a red chip.

Risky Event. After the auction and distribution of insurance, the experimenter reached into the bag of chips, stirred the chips noisily to reinforce beliefs of randomness, and drew a chip from the bag so that all subjects could see its color. Another experimenter entered the color of this chip at a control terminal so that the appropriate adjustments--$1 to all if a white chip was drawn and $4 loss to those without insurance if a red chip was drawn--could be made to the subjects’ balances and displayed on their terminals. To allow pooling of data across sessions and to ensure that all subjects received the same probabilistic experience, the drawing was controlled (the different colors of the chips were distinguishable by texture as in Phillips and Edwards, 1966, and many similar psychology experiments) according to the following sequences: [1]

[1] Subjects were chosen and sessions were arranged so that communication between subjects participating in different sessions outside of the laboratory was unlikely. In fact, the supposedly random draws were never questioned by subjects. Rather, subjects were suspicious that the computer run auction was rigged.

Experimental Sessions	Probability Level	Sequence of Chips (W= white R= Red)
Lower Probabilities	p = .01 p = .10 p = .20 p = .40	W W W W W W W W W W W W R W W W W W W W W W R W W W R W W W W R R W W W R W W W
Higher Probabilities	p = .60 p = .90	R W R R W W R W R R R R R W R R R R R R

Sequence. The different components of the experiment were presented and data were obtained in the following fixed order in the lower probability experimental sessions:

Inexperienced Hypothetical Bids at p = .2, .1, .01, and .4

Vickrey Auction Practice Bids, 4 rounds at p = .2

Vickrey Auction Binding Bids, 10 rounds at p = .2

Experienced Hypothetical Bids at p = .1

Vickrey Auction Binding Bids, 10 rounds at p = .1

Experienced Hypothetical Bids at p = .01

Vickrey Auction Binding Bids, 10 rounds at p = .01

Experienced Hypothetical Bids at p = .4.

Vickrey Auction Binding Bids, 10 rounds at p = .4.

In the higher probability experimental sessions the following fixed order was used:

Inexperienced Hypothetical Bids at p = .6 and .9

Vickrey Auction Practice Bids, 4 rounds at p = .6

Vickrey Auction Binding Bids, 10 rounds at p = .6

Experienced Hypothetical Bids at p = .9

Vickrey Auction Binding Bids, 10 rounds at p = .9

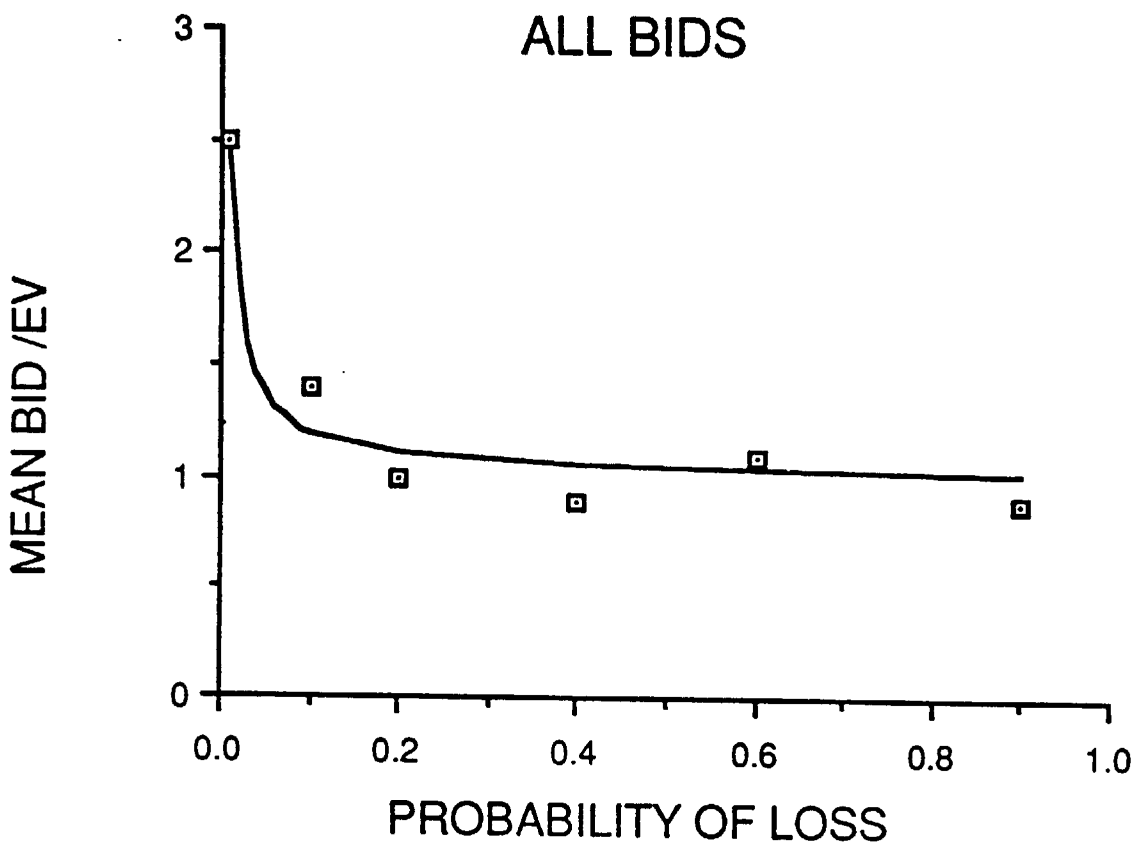
The fixed order of probabilities makes it impossible to have experienced hypothetical bids for p = .2 and p = .6 because these were always the first probability levels presented in the sequence of actual auctions.

3.3 Results

3.3.1 Overview

Summary statistics describing results of the experiment are presented in Figure 3.1. This figure depicts means of bids pooled across rounds divided by expected value, B/EV, plotted against probability of loss. As noted in Section 3.2.1, we normalize bids for insurance by dividing by expected values so we can directly compare results at different probability levels with each other and with the predictions of EUT. According to EUT we would, of course, expect mean measures of B/EV to equal unity. Note that, at probabilities of loss of .2 and above, mean B/EV is close to unity. However, at the lower probabilities of .1 and .01, EUT fails to predict observed values. The mean bid rises to about two and one-half times EV at a probability of loss of .01. Thus, on average, individuals overbid for insurance at low probabilities. This result at low probabilities is entirely consistent with the predictions of PT and can be interpreted as a direct consequence of the weighting function. From equation (3.2.10), PT predicts B/EV = $\pi(p)/p$ which should exceed unity for small p since it is assumed that $\pi(p)>p$ in this case. Mean auction values do not necessarily support PT at the higher probabilities (.2 and above) since PT argues that $\pi(p)< P$ for large P which implies B/EV<1. However, it should be noted that PT only predicts small underbidding at higher probabilities for the specific weighting functions typically proposed, so

FIGURE 3.1



we doubt that these data support a rejection of PT at higher probabilities. Rather, EUT and PT are similar in their predictions at higher probabilities for the case of insurance against loss.

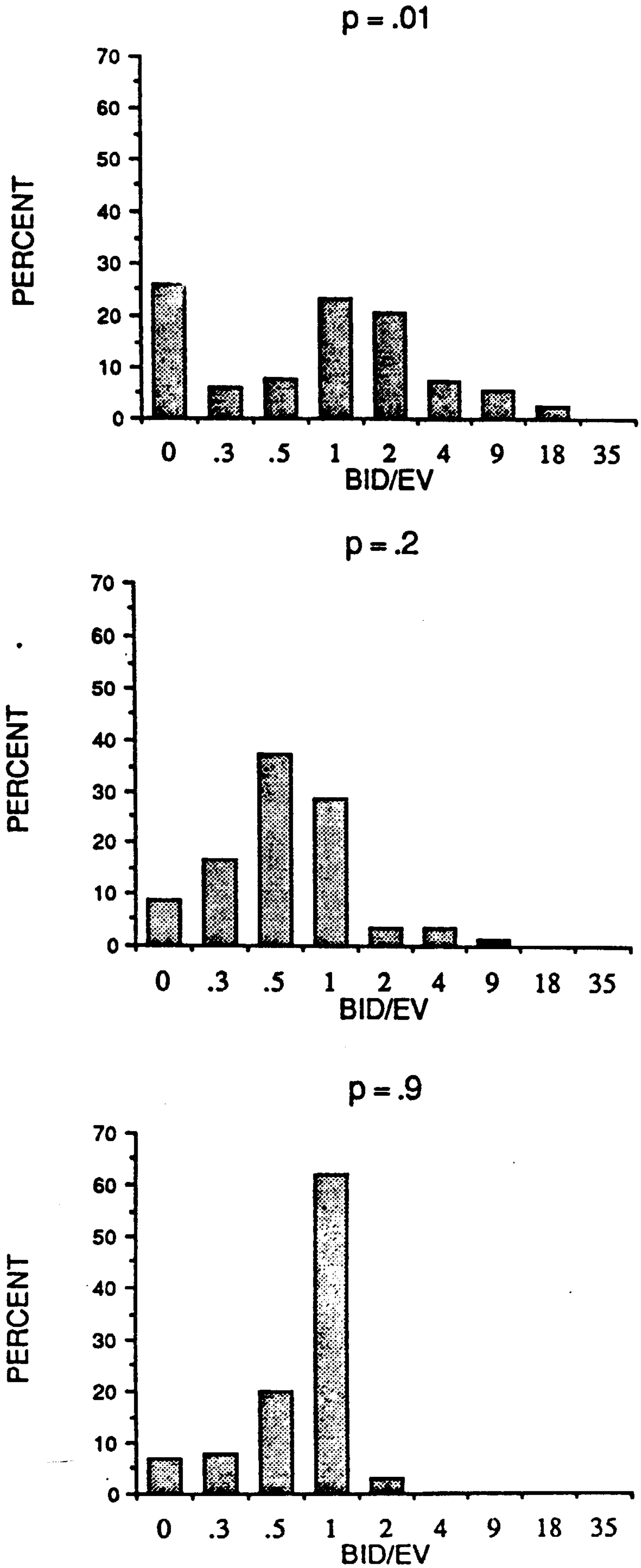
To attempt to understand the source of the large deviation from EUT that apparently occurs at low probabilities, we turn to a detailed analysis of the frequency distribution of B/EV. Figure 3.2 presents frequency distributions pooled across trials for auction values of B/EV at probabilities of loss of .9, .2 and .01. Since the frequency distributions for B/EV at probabilities of .1 and .4 are similar to that shown for .2, and since the distribution for .6 is similar to .9, only three distributions are presented. Also, since the variance increases greatly at lower probabilities a logarithmic horizontal axis is used to allow comparisons across probabilities. The approximate midpoint value of B/EV for each bin is shown under the bar representing the frequency of bids falling within the bin.[2]

The most striking feature in the top panel of Figure 3.2 is the pronounced bimodality of the distribution of bids which occurs at a probability of loss of .01. More than 25 percent of the bids in the sample are equal to zero, forming a lower mode. The distribution of positive bids on the logarithmic scale is approximately normal, thereby implying a log-normal distribution of the positive bids. The two modes suggest that two different cognitive processes may be operating at low probabilities.

---

[2] Bins were chosen as follows: The largest values of B/EV obtained in the experiment were equal to 50 and occurred at p=.01. A logarithmic scale was created by successive halving of this value. Thus, bins were created for values of B/EV <50 and >25, <25 and >12.5, ≤12.5 and >6.25, <6.25 and >3.125, <3.125 and >1.5625, <1.5625 and >.78125, ≤.78125 and >.390625, ≤.1953 and >zero. A separate bin was provided for zero bids. The rounded geometric means of the end points of each of the bins are shown along the horizontal axis of Figure 3.2.

FIGURE 3.2


p = .01
70
60
50
40
30
20
10
0
PERCENT
0
.3
.5
1
2
4
9
18
35
BID/EV
p = .2
70
60
50
40
30
20
10
0
PERCENT
0
.3
.5
1
2
4
9
18
35
BID/EV
p = .9
70
60
50
40
30
20
10
0
PERCENT
0
.3
.5
1
2
4
9
18
35
BID/EV

The first process, editing, occurs when individuals dismiss the risk and bid zero. The editing process is generally necessary since no decisionmaker can explicitly consider every possible risk, no matter how small, givdecision maker a finite time constraint on the decision process. Some rule of thumb or heuristic must be used to decide which risks are worth considering more carefully. Examination of the lower two panels in Figure 3.2 shows that the number of zero bids falls sharply as the probability of loss increases from .01 to .2 and .9. Thus, editing seems to depend on probability of loss in this experiment (where the loss is held constant).

The second process, anchoring and adjustment, attempts to explain the mental steps that individuals use to generate positive bids. The first step is, of course, the judgment not to edit, i.e., that the risk is worth considering. Second, individuals focus (anchor) on the loss, $4.00, and attempt to adjust the loss downward to account for the fact that the loss will occur only some of the time. Thus, for example, with a probability of loss of .1, subjects may be viewed as going through the following mental iterations:

<u>Example</u>

> “Should I bid $4.00? No, the loss will not occur all the time so insurance is not worth that much. Should I bid $2.00? No, this still seems to be too high a proportion of $4.00. Should I bid $1.00? Maybe. Should I bid $.50? Maybe. I think $.50 is probably closer than $1.00 to the proportion of $4.00 which represents the value of the risk of loss so I guess that will be my bid.”

Note that EV is $.40 in this case and, in the example above, the adjustment process has generated a bid which is quite appropriate. However, many subjects are likely to “guess" $1.00 since the intuitive process used in the example above is not highly accurate. That is, individuals may not engage in formal mathematical calculations in arriving at their bids. Further , since the process starts from the loss and, at least for the average individual, works downward, any error is likely to produce an upward bias in bids in that the adjustment process is likely to fall short as has been demonstrated in many studies of the anchoring and adjustment process. As is evident in examining the bottom panel in Figure 3.2, although the mean bid at p=.9 is near EV (B/EV = 1 in the figure) the variance is very large. Thus, some individuals adjust too far down while others mistakenly adjust upwards from the loss producing bids greater than the loss. Also, the distribution of bids is essentially normal (as opposed to log-normal) at p=.9, possibly reflecting more of a two way adjustment process either up or down from the loss anchor.

3.3.2 Models for Editing and Anchoring and Adjustment

We propose and test the following formal models to explain the data from this experiment. Define the fraction of zero bids as f° and the fraction of positive bids as $f^{+}=1-f^{0}$. Figure 3.2 suggests that f^{+} will be a function of p, $f^{+}(p)$. Figure 3.3 plots the fraction of positive bids versus probability of loss. The fitted curve shown in the figure labeled "model" is estimated using data pooled across trials for the six probabilities as

$$\text{(3.3.1)} \qquad f+ = \underset{(152)}{.936} - \underset{(13.2)}{.00}\frac{1}{p} .$$

$$DF = 4 \quad R^2 = .98$$

where t-statistics, testing whether the coefficients differ from zero, are shown in parentheses. Obviously the fraction of positive bids fails sharply as the probability falls and the amount of editing increases. The functional form used was chosen on the basis of fit. For example, replacing the l/p term with an exponential in p lowered the R^2 to a value near .4.

To model the anchoring and adjustment process, we focus on explaining

FIGURE 3.3

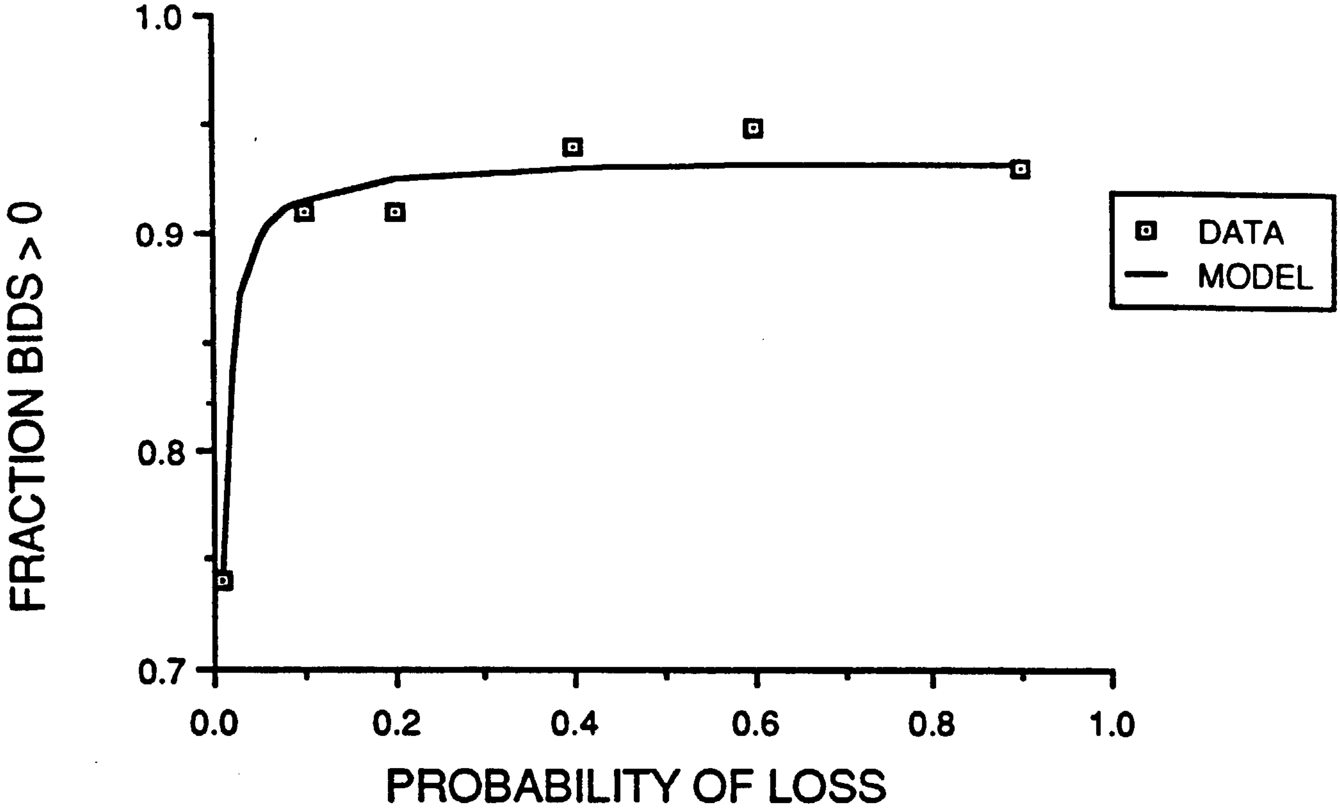
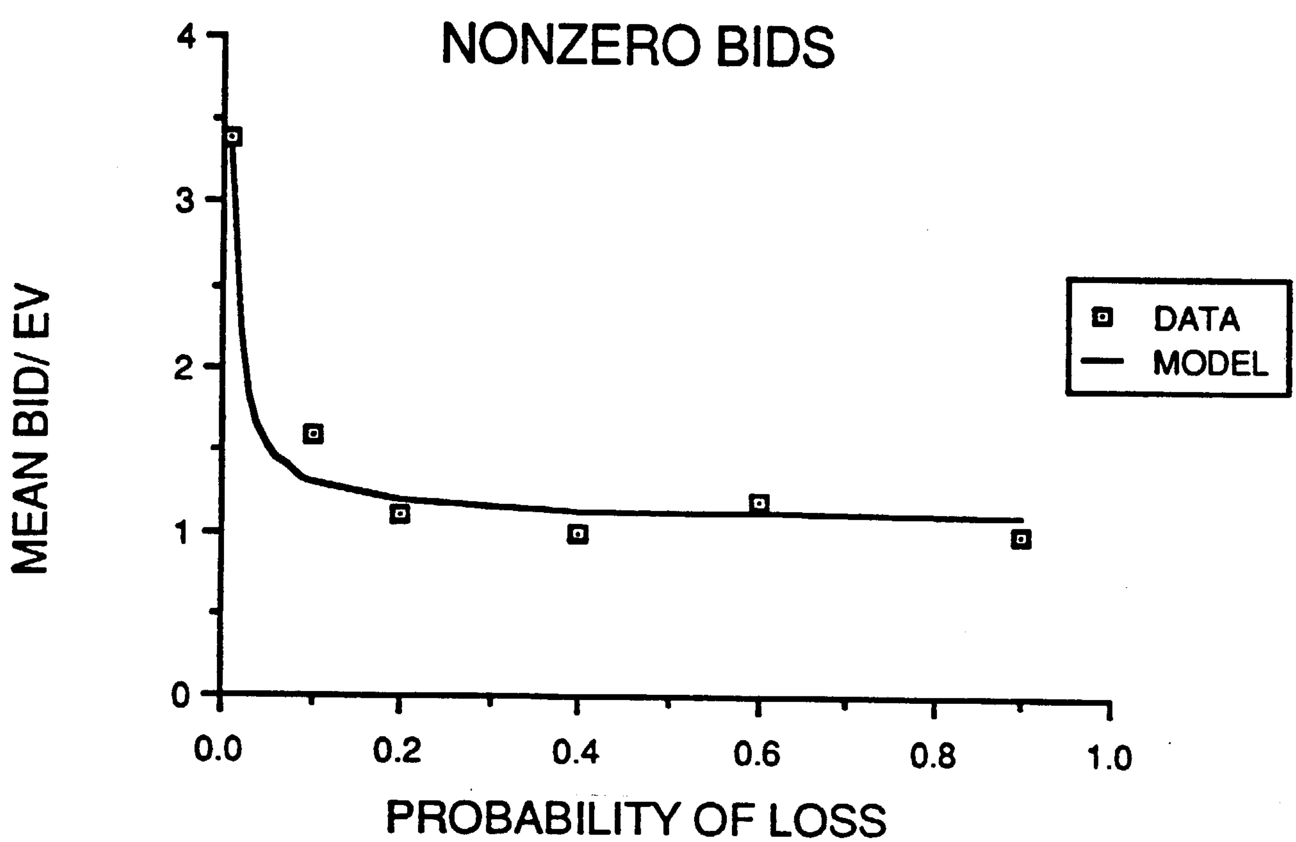


FIGURE 3.4



the mean value of positive bids, $\bar{B}^+$. The mean of all bids $\bar{B}$ is then equal to $f^+ \cdot \bar{B}^+$. Thus, the relationship developed for f^+ above which describes the editing process and the relationship developed below for $\bar{B}^+$ together yield the curve labeled "model" in Figure 3.1 that explains the average bid $\bar{B}$ in terms of p.

The average positive bidder is assumed to start from the loss, L, and move towards EV = pL. Thus, the distance over which the adjustment occurs is L-pL. Individual positive bids, B^+, can be viewed as being equal to EV plus an error term, ε, so

(3.3.2) $$B^+ = pL + \varepsilon.$$

The error term ε is assumed to have a distribution $g(\varepsilon, L-pL)$ which is shifted by the distance, L-pL, over which the adjustment process occurs. Further it is assumed that the mean of the error term, $\bar{\varepsilon}$ has the following properties

(3.3.3) $$\bar{\varepsilon}(L-pL) = \int_{-\infty}^{+\infty} g(\varepsilon, L-pL)\varepsilon d\varepsilon \begin{cases} =0 \text{ for } L-pL=0 \\ >0 \text{ for } L-pL>0 \end{cases}$$

so the greater the distance over which the adjustment must occur, the more the mean error exceeds zero.[3] We use a first order Taylor series linear approximation of (3.3.3) to obtain

(3.3*4) $$\bar{\varepsilon} \approx \bar{\varepsilon}(0) + \bar{\varepsilon}'(0) \cdot (L-pL)$$

where $\bar{\varepsilon}(0) = 0$ and $\bar{\varepsilon}'(0) > 0$ by (3.3.3). Thus, if we assume that adjustment is a linear process we can substitute (3.3.4) into (3.3.2) to obtain the

[3] We note that the anchoring and adjustment process can also possibly produce zero bids. Some fraction of the population may adjust to zero or beyond to a negative bid which presumably appears as a zero bid. The fraction of zero bids generated by anchoring and adjustment would, based on our model be

$$\int_{-\infty}^{-pL} g(\varepsilon, L-pL) d\varepsilon.$$

We do not attempt to account for this fraction in our statistical analysis, rather assigning all zero values as edits. The error introduced by this assumption is presumed to be small.

mean positive bid as:

$$(3.3.5) \qquad \bar{B}^{+} = pL + \bar{\varepsilon}'(0)\cdot(L-pL).$$

$\bar{\varepsilon}'(o)$ can be interpreted as a parameter which estimates the fraction of the distance L-pL that the average individual falls short in the attempt to adjust the bid from the anchor, L, to pL.[4] The anchoring and adjustment process is an example of a framing effect in which consideration of the loss as an anchor biases the estimation of EV. A non-random error has thus been introduced by the intuitive thought process used to estimate EV. It is the non-random nature of this type of error which makes analysis of cognitive processes of importance in understanding economic behavior in which low probabilities are involved.

To obtain a functional form for statistical estimation and testing of hypotheses we divide (3.3.5) by EV to obtain

$$(3.3.6) \qquad \frac{\bar{B}^{+}}{EV} = 1 + \bar{\varepsilon}'(o) \quad (\frac{1-p}{p}).$$

This relationship is estimated using the six observations on mean positive bids pooled across trials for the six alternative probabilities as

$$(3.3.7) \qquad \frac{\bar{B}^{+}}{EV} = \underset{(13.3)}{1.1} + \underset{(11.5)}{.023} (\frac{1-p}{p})$$
$$DF = 4 \qquad R^2 = .97$$

where the constant is free (not forced equal to unity). This estimated relationship is plotted in Figure 3.4 along with the data points. Using (3.3.7) we can test two hypotheses: The constant is not significantly different from unity (t(4)=1.2) but the coefficient on (1-p)/p is reliability different from, zero (t(4)=11.5). Thus, we can conclude that the probability level affects $\bar{B}^{+}/EV$.

---

[4] An anchoring and adjustment model which has some similarities to the one presented here has been successfully tested by Johnson and Schkade (1986) for a rather different experimental situation involving uncertainty. Their experiment did not involve the lower probabilities examined here and focused on having individuals provide hypothetical estimates of probability and certainty equivalents.

$\bar{B}^+/EV$ should, of course, simply equal unity and be unaffected by p if EUT holds. Thus, we must in general reject EUT. However, from (3.3.7), $(1-p)/p \to 0$ as $p \to 1$ and as a result $B^+/EV \to 1.1$, the constant, which is not significantly different from unity. Thus, we do not reject the hypothesis that EUT applies asymptotically for high probabilities; in Figure 3.4, $\bar{B}^+/EV$ is not very different from unity for p>.2. Finally, the data are consistent with the anchoring-and-adjustment model as expressed in (3.3.6) because the constant is similar to unity as predicted by the model and $\bar{\varepsilon}'(o)$ (estimated as .023) is significantly different from zero. This interpretation suggests that the average individual adjusts 97.7 percent of the distance from the loss to the expected value. The 2.3 percent shortfall in the adjustment process only leads to a large error in estimating EV (as a proportion of EV) at smaller probabilities as the distance between L and pL increases.

3.3.3 Comparison to other Models and Experiments

Returning to the original specification of PT we can now suggest an explanation for the weighting function $\pi(p)$, which can be interpreted as a decision weight on the loss L. The mean $\bar{B}/EV$ of the entire population from PT should be $\pi(p)/p$ from the analysis of section 3.2.1. Thus, using (3.3.6)

(3.3.8) $$\frac{\pi(p)}{p} = \frac{\bar{B}}{EV} = \frac{f^+\bar{B}^+}{EV} = f^+(p)\cdot(1+\bar{\varepsilon}'(o)\frac{1-p}{p}).$$

so

(3.3.9) $$\pi(p) = f^+(p)(p+\bar{\varepsilon}'(o)(1-p)).$$

The weighting function can thus be interpreted as an artifact of the editing process (described by the $f^+(p)$ relationship) and the anchoring-and-adjustment process (captured in the $\bar{\varepsilon}'(o)$ parameter). Unfortunately, at low probabilities where the divergence of $\pi(p)$ from p becomes important, the decision weight does not describe the behavior of any "typical" individual, but rather is the average of two divergent behaviors. In one behavior individuals bid zero. In the other, the mean individual bids well above EV, This bimodality has several implications for analyzing low-probability, high-loss events. For example, the intense conflict which often arises over technological risks such as those from nuclear power might be explained as a conflict between individuals from upper and lower modes similar to those apparent in the top panel of Figure 3.2. Such conflict cannot be explained in terms of the weighting function of PT which likely represents the average of a bimodal distribution. Rather, consistent with the spirit of PT and as an extension of PT we would propose the weighting function be discarded in favor of explicit modeling of the editing and the anchoring and adjustment processes. One way to formalize these notions is to refocus PT on the determinants of the fraction of positive bids, f^+, and on the determinants of the mean positive bid - expected value ratio, $\bar{B}^+/EV$.[5]

[5] Studies by Hershey and Shoemaker (1982), Shoemaker and Kunreuther (1979), and Slovic, et al. (1977) have also investigated insurance preferences. However, those studies are difficult to compare to the present study because those earlier studies did not involve market pressures, observed only hypothetical responses, and used very large losses (e.g., $100,000). The major difference in those studies is that respondents did not bid for insurance but only indicated whether they would accept or reject insurance offered at an actuarially fair price. In terms of our analysis, for lower probabilities these studies were essentially tracking f^+, the proportion of the sample in the upper mode. These studies are not consistent with each other in terms of their implicit modeling of f^+ as a function of p and L and it should be noted that we have offered only an atheoretical empirical model of f+ as a function of p. Hershey and Shoemaker suggest that some of the differences between studies are due to context or framing effects. We offer the further suggestion that comparisons between these studies will be difficult because the actuarially fair price will be dividing a bimodal distribution between modes for some combinations of p and L and a unimodel distribution (at approximately the mean) for other distributions. Clearly, much work remains to be done to understand the factors determining whether people edit or anchor-and-adjust from the loss.

3.3.4 Trial Dynamics

We now turn to an examination of trial dynamics - how experience over trials affects bids for insurance. Figure 3.5 shows mean auction values of B/EV (including zero bids) across rounds or trials. The means in Figure 3.5 remain constant and near unity across rounds for the higher probabilities of .2 and .9 shown (.4 and .6 are similar) but show a slight upward drift at .1 and a large upward movement at .01 across rounds. We interpret the upward drift over rounds of B/EV at the lower probabilities to be the result of gambler’s fallacy. That is, if a run of successive white chips is drawn, subjects become falsely convinced that the subjective probability of drawing a red chip has increased. This effect is not apparent at higher probabilities because when a red chip is drawn, subjects either “reset” their subjective probability close to the objective probability or assume that the odds of drawing another red chip have gone down. Thus, gambler’s fallacy appears to be self cancelling when subjects experience fairly frequent draws of a red chip. Of course at low probabilities, long runs of successive draws of white chips are likely and the cumulative effect of gambler’s fallacy will be apparent. When examining Figure 3.5, it is important to note that no red chips were drawn across the ten rounds at a probability level of .01. Also, at the probability level of .1 only one red was drawn (on the third round).

To analyze the mechanics of gambler’s fallacy we separate the data at P=.01 by again analyzing the fraction of positive bids and the mean B/EV of positive bids which are, on a round by round basis:

FIGURE 3.5

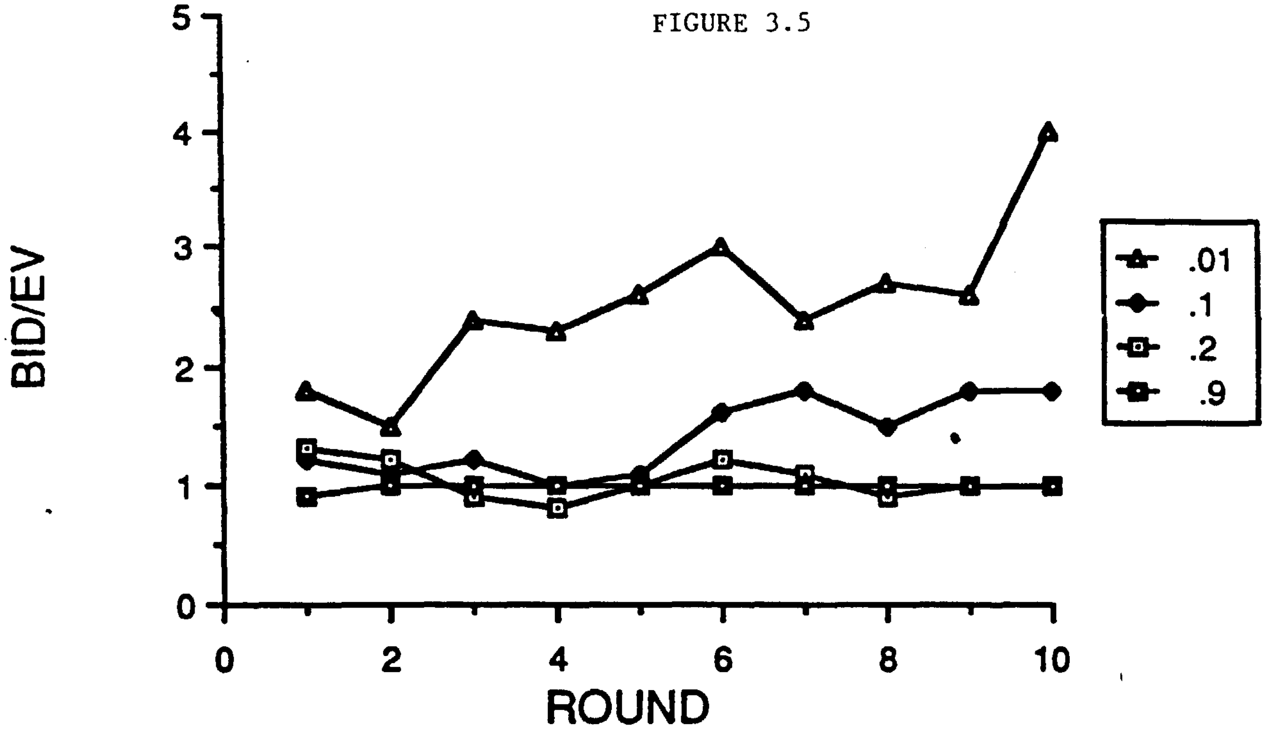
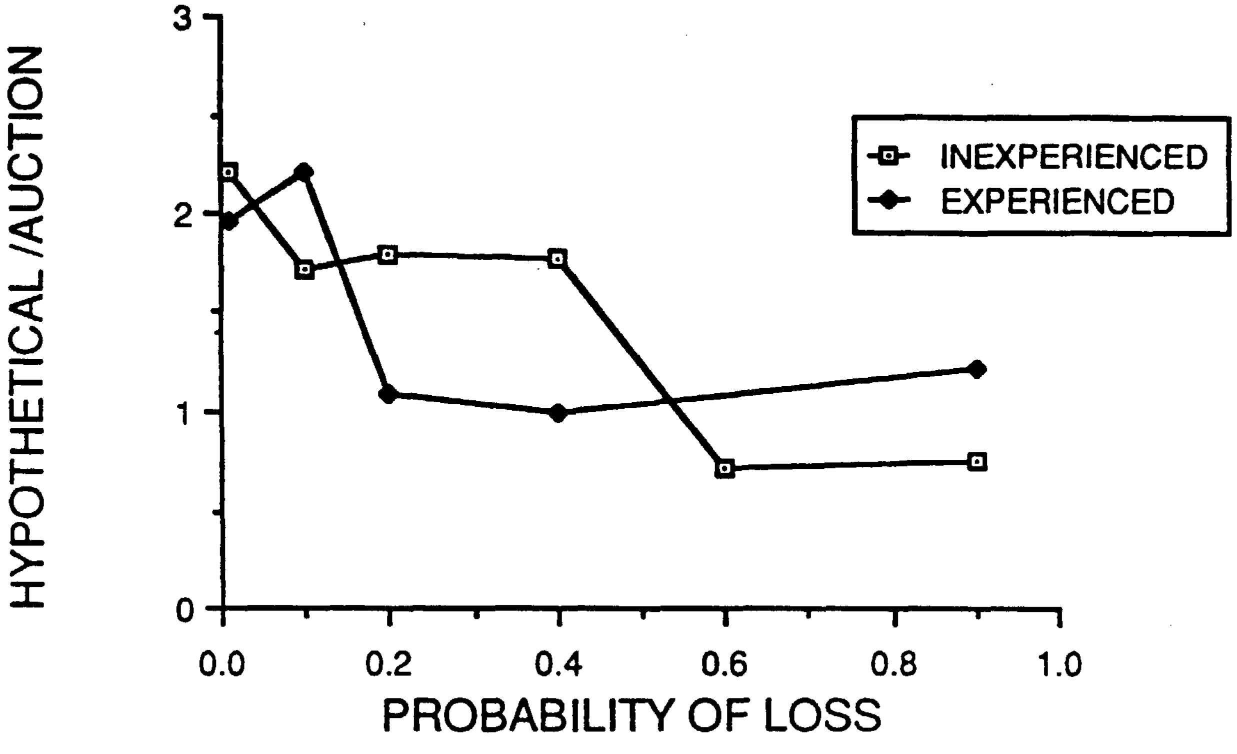


FIGURE 3.6



Round	1	2	3	4	5	6	7	8	9	10
f^+	.60	.62	.68	.75	.78	.78	.72	.85	.82	.80
$\bar{B}^+/EV$	3.0	2.3	3.5	3.0	3.4	3.9	3.3	3.2	3.2	5.0

These data suggest two observations. First, the fraction of positive bids shows a relatively steady increase across rounds. Thus, mode switching from the zero mode to a mode centered around B/EV=3.2 (the mean of B^+/EV for rounds 1 to 9) seems to be the source of most of the gambler's fallacy effect apparent in Figure 3.5. Second with the exception of the last round (which shows a strong end effect) $\bar{B}^+/EV$ is relatively stable across rounds. A linear regression of f^+ as a function of round yields an estimated equation (using only the first 9 rounds)

(3.3.10) $$f^+ = .59 + .028 \cdot \text{round}$$
(20.9) (5.6)

DF = 7 R^2 = .82

which contrasts sharply with a linear regression of $\bar{B}^+/EV$ on round,

(3.3.11) $$\bar{B}^+/EV = 2.89 + .065 \cdot \text{round}$$
(9.6) (1.2)

DF = 7 R^2 = .17

which has a coefficient not significantly different from zero for the round variable. However, we are not prepared to dismiss the hypothesis that successive rounds affect the mean positive" bid for two reasons.

First, the amount of adjustment which occurs may depend on experience; for example $\bar{\varepsilon}'(o)$ may decrease with more experience to produce bids close to the target pL. Since subjects had much experience at other probabilities prior to purchasing insurance at p = .01, the effect of experience on valuing a new risk may not be apparent here. Second, it is likely that gambler's fallacy affects the subjective probability belief of individuals. Both f^+ and B^+/EV seem to be functions of p. If we replace p with s, we should see an effect of gambler's fallacy on both f^+ and $\overline{B}^+/EV$. The next section describes an experiment at p=.01 in which subjects have no prior experience and in which the number of successive trials is raised to 50. This second experiment was specifically structured to further explore trial dynamics. In any case, the analysis above confirms another cognitive source of deviation from EUT, gambler's fallacy, which again, in a market like auction environment, seems to occur only as a problem at lower probabilities.

3.3.5 Hypothetical Behavior

As noted in the introduction of this chapter, psychology experiments of risky decision making have often used hypothetical bids and risks. In contrast, experimental economics traditionally employs actual financial transactions. The obvious question is whether using real monetary consequences differs from using hypothetical amounts. Figure 3.6 shows how means of hypothetical bids collected in our experiment compare to means of actual auction bids pooled across trials. Hypothetical mean bid divided by actual auction mean bid is shown on the vertical axis and probability of loss is shown on the horizontal axis. The inexperienced hypothetical bids collected at the start of the experiment clearly overestimate actual auction bids at low probabilities (since the ratio shown in Figure 3.6 is greater than one) and underestimate actual auction bids at high probabilities (since the ratio is less than one). The single deviation from the predictions of PT apparent in our auction results was that at high probabilities bids were close to EV, that is, we did not see the underweighting predicted by PT. We have no explanation as to why inexperienced hypothetical bids at high probabilities show underweighting and actual auction bids do not. In contrast, however, experienced hypothetical bids, which were collected after actual auction experience at other probabilities, were good predictors of auction bids at probabilities of .2 and above. It should be noted that the experienced hypothetical data point shown for .2 was not taken from the experiment described herein but from a pilot study where the order of probabilities was different so that an experienced hypothetical value could be obtained for p=.2.

Both inexperienced and experienced hypothetical bids are about twice actual auction bids at p=.1 and .01. We conjecture that the overestimation of hypothetical bids which occurs at low probabilities may be due to an incomplete adjustment process. In other words, since individuals start with the loss and work downward in deriving bids and since the distance between the loss and EV is great at low probabilities, practice may increase the amount of downward adjustment which occurs, bringing bids closer to EV. At the lower probabilities more adjustment is required and both inexperienced and experienced hypothetical bids may represent the first iteration in the adjustment process. In the experiment described in the next section subjects begin an actual auction at p=.01 with no prior laboratory experience of the auction procedure or this type of risk. If this hypothesis is correct, actual auction bids should start at very high values. From Figure 3.1, actual auction bids are about 2.5 times EV at p=.01. From Figure 3.6, hypothetical bids are about 2 times actual auction bids at p=.01. Thus, we conjecture that completely inexperienced actual auction bids might be 5 times EV. If this is the case, then hypothetical bids might be good predictors of completely inexperienced auction behavior. Note in this context that all of the auction behavior in the experiment described above was of the experienced type because we began the auctions with four non-binding practice trials.

3.4 A Laboratory Simulation of the Response to a ‘New” Risk

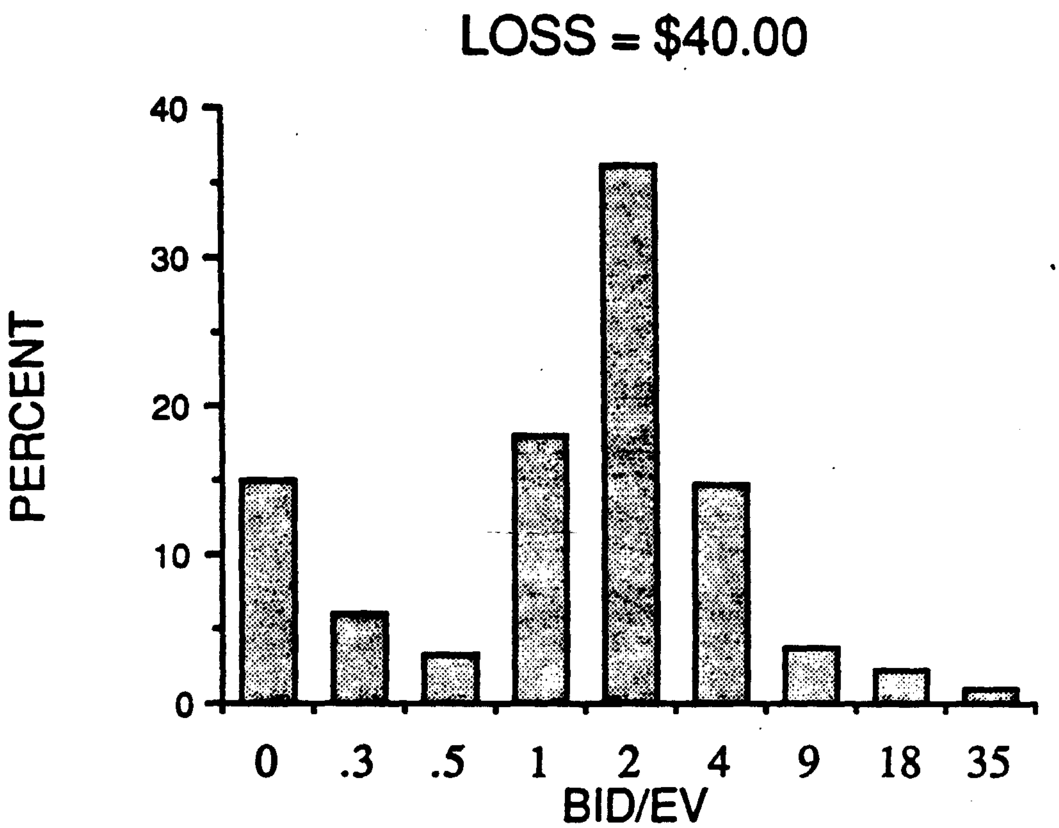
3.4.1 Overview

Given the bimodality and dynamic instability of values obtained at p=.01 in the experiment described above, a second experiment was conducted to explore further these phenomena at p=.01. A number of specific questions motivated the design of the new experiment. First, the editing phenomenon might have resulted from the relatively small $4 loss employed. Would editing still occur at similar frequencies for a much higher loss? Will bimodality still characterize the distribution of bids? To address such questions, we raised the loss to $40. Second, subjects in the experiment described above faced odds of loss of .01 after they had obtained a considerable amount of experience both with the Vickrey auction for insurance and with other probabilities. Many real world policy problems are associated with the response to new risks. Populations are informed that the landfill near which they have lived for a decade is leaking toxic substances or that the long dormant volcano or fault now poses a threat. New technologies are often resisted because they are viewed as dangerous, but later become accepted. Thus, to explore these issues, the auction was begun with no practice trials and the number of rounds was increased to 50 so that the lengthy recurrence intervals between events (draws of a red chip) characteristic of real world low probability hazards could be simulated in the laboratory. A red chip was drawn on round 33 so that subjects could accumulate benign experience in the form of a lengthy sequence of white chips before the event occurred. Seventeen remaining rounds were then available to examine behavior after the event occurred. As in the previous experiment, each of the six sessions employed eight subjects drawn from undergraduate economics classes, a Vickrey auction was used to sell four insurance policies in each round, and each subject was given one dollar if a white chip was drawn to help finance successive trials. The risky situation, initial balance and Vickrey auction were described to subjects both in written instructions and in an oral explanation which allowed questions. Subjects were shown 99 white chips and 1 red chip as they were placed in a bag. The sequence of 50 binding actual auctions then began immediately. A chip was drawn and replaced following each auction.

3.4.2 Results

Figure 3.7 shows the frequency distribution of B/EV pooled across all 50 trials. Generally, the distribution of B/EV for the $40 loss looks remarkably similar to the frequency distribution shown for the $4 loss in the top panel of Figure 3.2. Both distributions are strongly bimodal with

FIGURE 3.7



one mode at zero and another above EV (EV is shown as B/EV=1 in the figures). Since the horizontal axis is logarithmic, the upper modes in both cases appear to center on approximately log-normal distributions. Two minor differences are also apparent. First the upper mode for the $40 loss is shifted slightly to the right compared to the $4 loss. As we show below when we examine trial dynamics, inexperienced bids for the case of the $40 loss were very high in the early rounds. Thus, the difference in the initial amount of experience between the two experiments likely explains this shift. Second, in the $40 loss experiment some bidders seem to be adjusting upwards from a zero anchor creating a descending step pattern (moving to the right) for the zero, .3 and .5 B/EV bins in Figure 3.7. This suggests that individuals who edit are in reality choosing a zero anchor as opposed to the loss anchor as the basis for an upward as opposed to downward adjustment process. This leads to the conjecture that a lower mode just above zero, made up of individuals who edit might evolve under some circumstances.[6] In any case, the pronounced bimodality of the

[6] To understand these circumstances we need to consider why we do not see evidence of upward adjustment in the case of the $4 loss shown in the top panel of Figure 3.2. Subjects demonstrated a strong tendency to submit bids in round monetary values such as $.00, $.05, $.10, $.25, $.50, $.75, $1.00, $1.50, $2.00, $3.00, $5.00, $10.00 and so on. This monetary anchoring has often been tested in psychological studies of decision making (e.g., Combs, Bezenbinder, & Good, 1967) and in survey research. Since EV was $.04 in the $4 loss experiment, no strong monetary anchors fell between $.05 and $.00 so any upward adjustment from zero likely fell in the B/N=1 bin. In contrast, with a $40 loss, EV=$.40, and monetary anchors of $.10 and $.25 were available in the B/EV=.3 and the B/EV=.5 bins respectively. Thus, the larger the loss, the more apparent upward adjustment from the zero. anchor will be in experiments of this sort. For very low probabilities and very large losses a mode just above zero might then become apparent since monetary anchors will be available just above zero.

earlier experiment is present at the higher loss, consistent with the editing and the anchoring and adjustment models developed in the previous section.

Trial dynamics are shown in Figures 3.8 and 3.9. Figure 3.8 shows mean bid divided by EV. Early bids for insurance averaged about five times expected value, which is consistent with our conjecture of the previous section, but decreased to about two times expected value just before a red chip was drawn on round 33. Reigning price (shown in Figure 3.9), after an initial rise, remained constant at about 2 1/2 times EV until, following the draw of the red chip, a sharp drop in reigning price occurred in round 34. Both mean bid and reigning price then increased to the conclusion of the experiment at round 50. We conjecture that bids fell in early rounds both because individuals gained experience (i.e., learned to adjust more completely) and because benign experience may work in the opposite direction from gambler’s fallacy by reducing the subjective probability of loss. Note that in the 10 round experiment subjects both had experience in forming values at other probabilities prior to facing odds of loss of .01 and had actually experienced the loss of $4 on the draw of a red chip. After the draw of the red chip in the $40 loss experiment bids also rose over following rounds as in the $4 loss experiment. Thus, we suspect that, in the absence of the experience of loss, draws of white chips may convince some that they should dismiss the risk and bid lower or bid zero for insurance. Experience with loss, however, seems to reverse this process. Convinced by experience that the loss can occur, some subjects seemingly felt that successive draws of white increased the need for insurance. The actual odds of drawing red remain constant over trials since the drawn chip

FIGURE 3.8

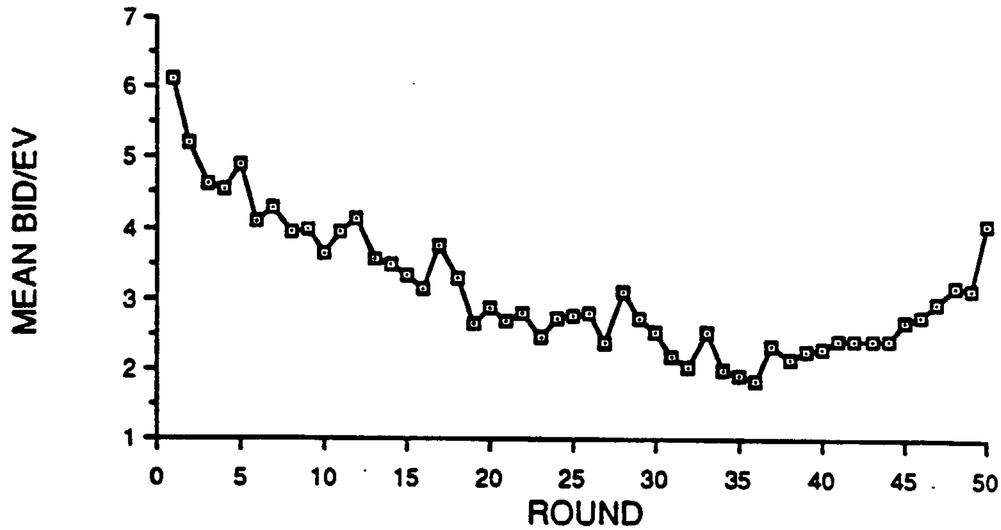
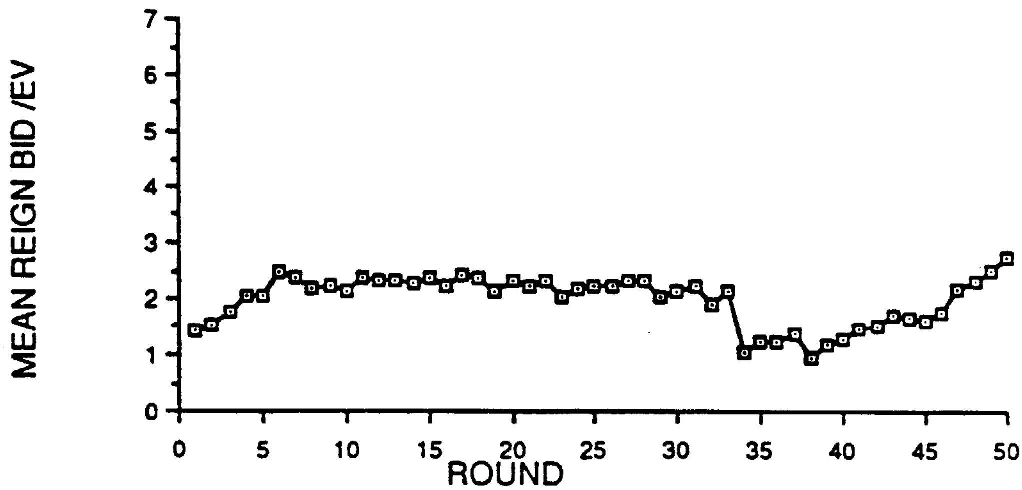


FIGURE 3.9



was replaced before the next trial. Thus, the possible benign experience and the gambler's fallacy effects on subjective probability of loss represent simple logical errors concerning probability.

To test these hypotheses we again split the data into two portions. One portion consists of the edit bids and the other contains all those bids presumed to be generated by anchoring on the loss L = \$40 and then adjusting downwards towards pL. This time we count as edit bids not only the zero bids but also those bids that are slightly above zero but still in the lower mode of the frequency distribution of Figure 3.7. The category for B/EV = .5 appears to be the boundary between the two groups of bids so we use the midpoint of that category as the dividing line. The fraction of bids in the upper portion is now f^+ and the mean of the bids divided by expected value in the upper portion is B^+/EV. Both f^+ and B^+/EV are calculated for each round and plotted in Figures 3.10 and 3.11 respectively. To develop a statistical model of the effects of benign experience and gambler's fallacy on f^+ and $\overline{B}^+/EV$ we define the following variables on the basis of round:

	Round
	1, 2, ..., 33, 34, 35, 36, ..., 50
Variable	
Benign	-33, -32, ..., -1, 0, 0, 0, ..., 0
After Red	0, 0, ..., 0, 1, 1, 1, ..., 1
Since Red	0, 0, ..., 0, 0, 1, 2, ..., 16

---

[7] If the prior definition of edit bids as only zero bids is used, the functional form of the analysis which follows is essentially unchanged. The only difference is that f+is much less stable when editing is restricted to zero bids. This suggests that those editing sometimes switch between zero and very low bids in the .3 and .5 B/EV categories of Figure 3.7.

FIGURE 3.10

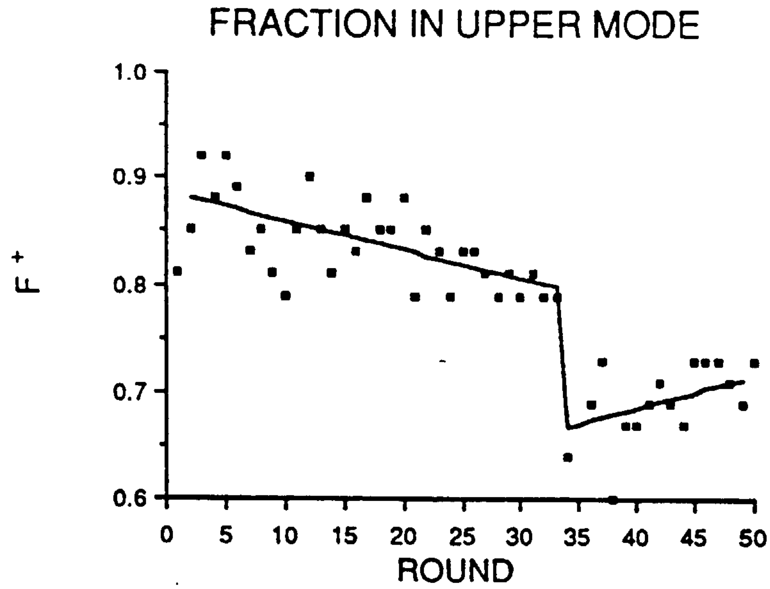
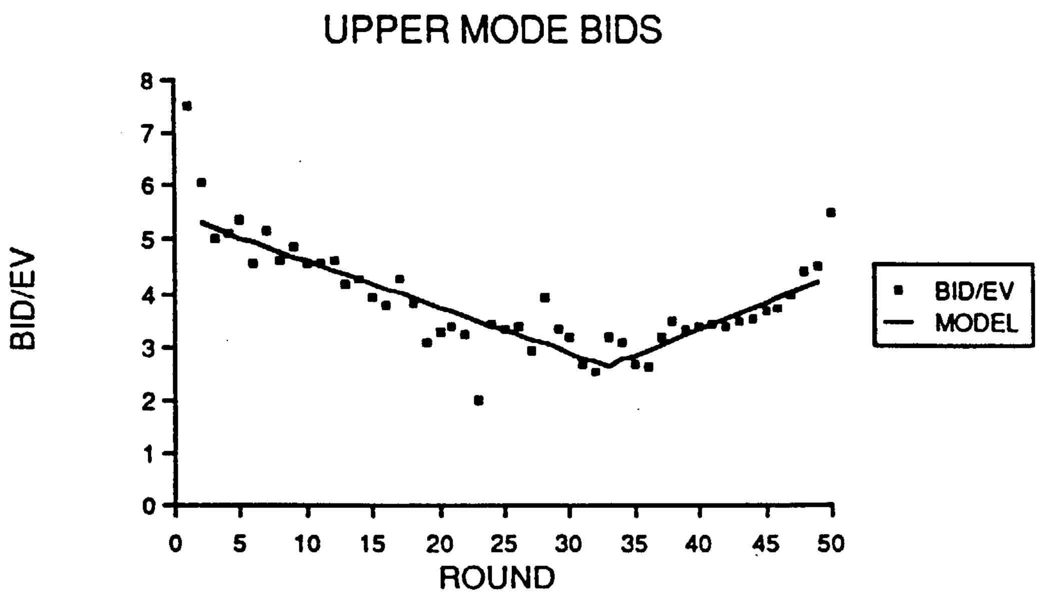


FIGURE 3.11



Linear regression estimates (excluding data from the first and last rounds, which are obvious obvious from Figure 3.11) are:

$$(3.4.1)\quad \underset{(73)}{f^{+}=.793}\;\underset{(-4.6)}{-.0027}\ (\text{Benign})\ \underset{(-7.1)}{-\ .128}\ (\text{After Red})\ \underset{(1.1)}{+\ .003}\ (\text{Since Red})$$

$$DF = 44 \qquad R^2 = .87$$

and

$$(3.4.2)\quad \frac{\overline{B}^{+}}{EV} = \underset{(22)}{2.54}\;\underset{(-14)}{-.086}\ (\text{Benign})\ \underset{(-1.1)}{+\ .21}\ (\text{After Red})\ \underset{(5.6)}{+\ .099}\ (\text{Since Red})$$

$$DF = 44 \qquad R^2 = .88$$

We can interpret the data shown in Figures 3.10 and 3.11 using the statistical model above as follows: The fraction of bids in the upper mode (those presumed to be adjustments from the anchor L=$40), f^{+}, begins (after only one round of prior experience) at approximately .88 and decreases significantly with benign experience (t(44)= -4.6) to approximately .80 just before a red chip is drawn on round 33. The draw of a red chip causes a sudden and statistically significant (t(44) = 7.1) drop in f+ (measured by the “After Red” intercept shifter) in round 34 to approximately .67. Gambler’s fallacy then appears to explain the increase in f^{+} to .71 by the next to the last round of the experiment but the “Since Red” variable does not have statistical significance (t(44) = 1.1).

The statistical model for the mean upper mode bid divided by expected value indicates that $\overline{B}^{+}/EV$ begins at approximately 5.3 and falls steadily to about 2.5 just before a red chip is drawn on round 33. This decline, associated with benign experience, is significant (t(44) = -14) and is possibly caused by an improvement in the adjustment process which occurs with experience or by a negative effect of successive draws of a white chip on the subjective probability of loss. In contrast to the f+ relationship, the draw of a red chip had a negligible immediate impact on $\overline{B}^{+}/EV$ as measured by the "After Red" intercept shifter (t(44) = -1.1). This difference supports a model of the decision structure which separates the editing decision from the process used to derive a positive bid. Note in this context that the drop in reigning price which follows the draw of a red chip on round 33 as shown in Figure 3.8 is the result of the drop in f^{+} which reflects a sharp increase in editing due to a gambler's fallacy effect (a draw of red reduces the chance of red on the next round). Positive bids, as measured by $\overline{B}^{+}/EV$, are however affected by successive draws of white after a red in a manner consistent with gambler's fallacy. As measured by the "Since Red" variable, gambler's fallacy is significant (t(44) = 5.6) in apparently increasing the subjective probability beliefs of positive bidders so that B^{+}/EV climbs to approximately 4.2 by the next to last round. Gambler's fallacy could be motivated in part by an end effect in that subjects increasingly attempt to defend their balances as the last round approaches.

The dynamic pattern shown in Figures 3.8-3.11 is broadly consistent with the conventional wisdom concerning subjective risk beliefs about natural and man-made hazards. Such risks are viewed as being overestimated both when people first become aware of the possibility of a catastrophe as well as in some period following the occurrence of a catastrophe. Risk beliefs are viewed as being underestimated following long periods of benign experience that inevitably occur given the long recurrence intervals of low probability events. We were not able to demonstrate fully this last supposed attribute because, although $\overline{B}$/EV declined steadily with benign experience, its value did not fall below unity nor did f^+ approach zero in 33 rounds. Obviously, the asymptotic properties of both f^+ and $\overline{B}^+$/EV need to be explored in future experiments involving more trials.

The bimodality present in these experiments suggests that actual insurance markets for disasters such as floods, earthquakes, etc. are likely to be peculiar. If insurance is provided by competitive suppliers, in the long run the offered price of insurance should be equal to expected value of the loss plus a minimum of administrative and transaction costs. Thus, insurance should be offered just above expected value. For low probability hazards, the offered price of insurance is likely to fall between the upper and lower modes of the frequency distribution of bids for disaster insurance. Thus, nearly all of the individuals in the upper mode (who anchor on the loss) will likely purchase insurance, while all of the individuals in the lower mode (who edit) are not likely to purchase insurance. In other words, the editing decision will completely control the number of insurance policies sold. Since benign experience appears to decrease f^+, which can be interpreted as the fraction of the population at risk likely to buy insurance, people may be reluctant to insure against disasters which have not occurred recently. On the other hand, sales of such insurance might well increase in a period following an event.

3.5 Can Risk Aversion Account for the Results?

Intuition suggests that the relative impact of risk aversion on bids for insurance should increase as the probability falls. Thus, in our first experiment with a loss of $4, one would expect little evidence of risk aversion at a probability of .9 and our results do show mean bids equal to expected value. However, at odds of .01 of a $4 loss, enormous risk aversion <u>could</u> raise mean bids to 2.5 times expected value as our data indicate. This possibility is excluded by the results from our second experiment with a loss of $40 as follows: To assume risk aversion is sufficient to explain our results for a $4 loss in wealth implies a very highly curved utility function in the neighborhood of the current level of wealth. If the size of the loss is increased from the $4 level to $40 level, as is done in our second experiment for odds of loss of .01, then given such a highly curved utility function, the effect of risk aversion should be to dramatically increase the ratio of bids to expected value above the 2.5 obtained in the $4 loss experiment. This does not happen. Rather, Figure 3.7 looks almost identical in pattern to the top panel of Figure 3.2. In other words, the frequency distribution of bids divided by expected value is about the same in the $4 loss case as it is in the $40 loss case. This is very strong evidence that relative risk aversion plays almost no role in our experiments. The enormous degree of risk aversion necessary to use expected utility theory to explain the 2.5 ratio of mean bid to expected value obtained at a .01 probability with a $4 loss implies that the frequency distribution shown in Figure 3.7 ($40 loss) should be far to the right compared to the frequency distribution shown in the top panel of Figure 3.2 ($4 loss). In fact, the slight rightward shift shown in Figure 3.7 compared to Figure 3.2 is much more likely attributable to the fact that we began our second $40 loss experiment with no prior auction experience for subjects to simulate what would happen when people faced a new risk. In contrast, subjects facing the $4 loss at .01 odds had prior experience with 4 practice trials plus 10 binding trials at odds of .2 plus 10 binding trials at odds of .1. In other words, learning, not risk aversion, likely explains any increase in bids relative to expected value in the $40 versus $4 loss experiments (both at .01 odds).

A second demonstration of the inability of risk aversion to explain our results comes from empirical studies of the coefficient of relative risk aversion. Following the notation used above, utility is U(Y) where Y is wealth. If we assume risk aversion, then $U''<0$. The coefficient of relative risk aversion is defined as $c=(-U''/U')\cdot Y$, a positive number. The empirical evidence on the coefficient of relative risk aversion has recently been summarized as follows:

> In particular, Cohn et al. (1975) found evidence that the coefficient of relative risk aversion is slightly decreasing in wealth. Friend and Blume (1975) found that "if there is any tendency for increasing or decreasing proportional risk aversion, the tendency is so slight that for many purposes the assumption of constant proportional risk aversion is not a bad first approximation" (p. 915). More recently, Morin and Suarez (1983) found the coefficient to be slightly decreasing for wealth levels up to $100,000, after which it becomes approximately constant. Furthermore, Friend and Blume estimated the market price of risk to determine a value for the coefficient, which they argue is greater than one and may be as high as two.[8]

In contrast to the field studies cited above, what value for c is implied by our experiments if risk aversion is to account for the observed increase in mean bid for insurance relative to expected value at a probability of loss of .01? Using the same notation as before, where we defined

Y° = initial wealth,

P = odds of loss (red chip),

L = size of monetary loss,

B = bid for insurance against loss,

(1-p) = odds of gain (white chip),

and G = size of monetary gain,

[8] "A Test of the Expected Utility Model: evidence from Earthquake Risks," by Brookshire, Thayer, Tschirhart and Schulze, JPE, 1985 Vol. 93:2, p. 381.

the true bid for insurance can be obtained in an expected utility framework by setting the expected utility of paying and obtaining insurance for \$B equal to expected utility without insurance:

$$p\,U(Y^\circ-B)+(1-p)U(Y^\circ+G-B)=pU(Y^\circ-L)+(1-p)U(Y^\circ+G).$$

To incorporate the coefficient of relative risk aversion, we substitute a second order Taylor series approximation of U(Y) into the expression above and obtain (where expected value = EV=p·L) an expression for bid divided by expected value (B/EV):

$$B/EV=1+\frac{c}{Y^\circ}\left[\frac{1}{2}(1-p(B/EV)^2)L+(1-p)(B/EV)G\right].$$

Note, that if no risk aversion is present, c=0, and B/EV=1. This is the assumption used in our prior analysis and, for losses of \$4 and \$40, can be justified as follows: The largest known value from field studies for c is about 2. In these studies, Y° = \$100,000. Thus, the term in square brackets above is multiplied by 2/\$100,000. In our \$40 loss experiment at odds of .01 (where the effect of risk aversion should be greatest in our experiments) the relevant observations on p,L, G and B/EV are .01, \$40, \$1, and 2.5 respectively. Using these values in the term in brackets above yields a value for that term of about 21.2. Multiplying c/Y° times this in the formula above gives a B/EV of 1.0004. Obviously this value of B/EV is inconsistent with our experimental results and with our use of B/EV=2.5 in the r.h.s. of the formula above. The quadratic formula can be used to solve for B/EV assuming c=2 and gives values of B/EV negligibly different from unity. Clearly, existing field evidence on risk aversion justifies our assumption of risk neutrality for losses of \$4 and \$40 used in our experiments.

Another approach is to ask what could c have to be to explain our results.

The expression above can readily be solved for c in terms of Y° B/EV, p, L and G. Using data from our $4 loss experiment at .01 odds gives c=33,333 (assuming Y°=$100,000, B/EV=2.5, p=.01 L=$4 and G=$1). For the $40 loss experiment c=6,667 (assuming Y°=$100,000, B/EV=2.5, p=.01, L=$40 and G=$1). Thus, not only must risk aversion take on absurd levels, but the degree of risk aversion must be much larger for small losses than for large losses. As noted above, field studies of risk aversion have shown c to be relatively constant for large changes in wealth. In summary the empirical evidence on the coefficient of relative risk aversion suggests that risk aversion plays no role in our experiments. This is consistent with our earlier argument that the increase in loss from $4 to $40, did not shift the frequency distribution of B/EV to the right.

3.6 Conclusion

The principal objective of the experiments reported in this chapter was to explore insurance behavior in a laboratory market-like environment where the probability of loss was varied. Thus, the predictions of expected utility theory as well as models from psychology could be compared against actual behavior at both higher and lower probabilities of loss. Additionally, repeated trials were included in the experiments so that the effect of experience on decision making could be determined. The results of the experiments suggest that although expected utility theory is an adequate explanation of behavior at higher probabilities of loss, at lower probabilities a much more complex model is required to explain observed behavior. This complex model has been evolving within psychology principally under the guise of prospect theory and includes features such as the editing phenomenon and the anchoring and adjustment process documented in our results. Additional results of our experiment include, at low probabilities, a large gambler’s fallacy effect and strong bimodality. These results are consistent with the direction and spirit of prospect theory. Further, they serve to reinforce our general conclusion that models which arise from psychology and which consequently focus on the mental processes and possible errors in those processes are central to any explanation of economic behavior motivated by low probability events.

Although it can be argued that markets themselves seem to promote behavior consistent with expected utility, they do not seem to help very much for low probability, uncertain situations, at least within the range of experience observed in our experiments. This implies that decision making at low probabilities is likely to be subject to error even in a market context. Individual responses to threats from low probability hazards such as Superfund sites are likely to suffer from the entire litany of cognitive difficulties identified above.

Although behavior differs from predictions of expected utility theory due to these cognitive difficulties, it is not appropriate to call behavior at low probabilities irrational. Anyone attempting to consider seriously the myriad of low-probability natural and technological hazards would quickly be overwhelmed and paralyzed with indecision. Therefore, it is rational to edit away many hazards that appear to be unlikely and to concentrate only on those that appear to be somewhat more likely. For those risks that are worth considering, the anchoring and adjustment process may produce estimates that are in many cases “close enough” in the sense that additional cognitive effort would not generally be worth its cost. This viewpoint is similar to that expressed in the recent paper by Russell and Thaler (1985). However, for very low probabilities and very high losses the intuitive reasoning that leads to the bimodality shown in our results has important implications for public policy.

For example, consider a controversy about whether a landfill containing toxic materials needs to be cleaned up when the scientific estimate of the risk is low. An application of the results from this study would lead us to expect that some residents living near the landfill would edit and therefore dismiss the risk. Other residents would consider their response to the situation by anchoring on the losses, which could be extreme such as cancer and birth defects, and then adjusting downward. Given such extreme anchors, the judged levels after insufficient adjustment, even if the percentage misadjustment factor is small, are likely to be quite high. This produces two groups of residents who disagree strongly about what needs to be done. One group complains that the risk is negligible and that all the fuss will only lower property values while the other group cannot understand why the former group is not concerned about the deadly risk confronting them all.[9] Chapter 4 shows that a drop in property values near a hazardous waste site seems to be associated with a bimodal distribution of risk beliefs very similar to that shown in the laboratory results presented here.

---

9 Note that we obtained this bimodality in the laboratory where subjects were able to perceive the risk directly by actually viewing the number of red and white chips put into the bag. For a real risk where such direct perception of the risk is not possible, there is likely to be wide individual variation in the estimation of the risk probability which will likely exacerbate the difficulties.